U.S. DEPARTMENT OF COMMERCE
National Technical Information Service
PB80-133119

Computer Program for Reducing Mass Spectra Data from GC/MS Systems

Research Triangle Inst, Research Triangle Park, NC

Prepared for

Environmental Sciences Research Lab, Research Triangle Park, NC

Oct 79

TECHNICAL REPORT DATA

(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA-600/2-79-194	2.	3. RECIPIENT'S ACCESSION NO. PB80-133119
4. TITLE AND SUBTITLE COMPUTER PROGRAM FOR REDUCING MASS SPECTRA DATA FROM GC/MS SYSTEMS		5. REPORT DATE October 1979
		6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) D. Rosenthal		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Research Triangle Institute P.O. Box 12194 Research Triangle Park, N.C. 27709		10. PROGRAM ELEMENT NO. 1AD712 BD-014 (FY - 78)
		11. CONTRACT/GRANT NO. 68-02-2745
12. SPONSORING AGENCY NAME AND ADDRESS Environmental Sciences Research Laboratory - RTP,NC Office of Research and Development U.S. Environmental Protection Agency Research Triangle Park, N.C. 27711		13. TYPE OF REPORT AND PERIOD COVERED Final 9/77 - 3/79
		14. SPONSORING AGENCY CODE EPA/600/09
15. SUPPLEMENTARY NOTES		

16. ABSTRACT

Existing computer programs for deconvoluting mass spectrometry data derived from gas chromatography/mass spectrometry/computer systems were evaluated, and the most efficient method was introduced unto an EPA UNIVAC computer. The program chosen was that reported by Rindfleisch and his co-workers at Stanford University. The accompanying program listing is included. The Rindfleisch program was tested with both standard runs and environmental samples. Sample components were successfully identified by an independent computer search program with an accuracy approaching 80%.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Air pollution		13B
*Computer system programs		09B
*Data reduction		20H
*Mass spectra		07D
Gas chromatography		
Mass spectroscopy		

18. DISTRIBUTION STATEMENT RELEASE TO PUBLIC	19. SECURITY CLASS *(This Report)* UNCLASSIFIED	21. NO. OF PAGES 171
	20. SECURITY CLASS *(This page)* UNCLASSIFIED	22. PRICE A09 - A01

EPA Form 2220-1 (9-73)

United States
Environmental Protection
Agency

Environmental Sciences Research
Laboratory
Research Triangle Park NC 27711

EPA-600/2-79-194
October 1979

Research and Development

PB80-133119

EPA

Computer Program for Reducing Mass Spectra Data from GC/MS Systems

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

RESEARCH REPORTING SERIES

Research reports of the Office of Research and Development, U.S. Environmental Protection Agency, have been grouped into nine series. These nine broad categories were established to facilitate further development and application of environmental technology. Elimination of traditional grouping was consciously planned to foster technology transfer and a maximum interface in related fields. The nine series are:

1. Environmental Health Effects Research
2. Environmental Protection Technology
3. Ecological Research
4. Environmental Monitoring
5. Socioeconomic Environmental Studies
6. Scientific and Technical Assessment Reports (STAR)
7. Interagency Energy-Environment Research and Development
8. "Special" Reports
9. Miscellaneous Reports

This report has been assigned to the ENVIRONMENTAL PROTECTION TECHNOLOGY series. This series describes research performed to develop and demonstrate instrumentation, equipment, and methodology to repair or prevent environmental degradation from point and non-point sources of pollution. This work provides the new or improved technology required for the control and treatment of pollution sources to meet environmental quality standards.

This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.

EPA-600/2-79-194
October 1979

COMPUTER PROGRAM FOR REDUCING MASS SPECTRAL DATA FROM GC/MS SYSTEMS

by

David Rosenthal
Research Triangle Institute
Post Office Box 12194
Research Triangle Park, North Carolina 27709

Contract No. 68-02-2745

Project Officer

Kenneth Krost
Atmospheric Chemistry and Physics Division
Environmental Sciences Research Laboratory
Research Triangle Park, North Carolina 27711

ENVIRONMENTAL SCIENCES RESEARCH LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
RESEARCH TRIANGLE PARK, NORTH CAROLINA 27711

iA

DISCLAIMER

This report has been reviewed by the Environmental Sciences Research Laboratory, U.S. Environmental Protection Agency, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the U. S. Environmental Protection Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

ABSTRACT

Existing computer programs for deconvoluting mass spectrometry data derived from gas chromatography/mass spectrometry/computer systems were evaluated, and the most efficient method was introduced unto an EPA UNIVAC computer. The program chosen was that reported by Rindfleisch and his co-workers at Stanford University. The accompanying program listing is included. The Rindfleisch program was tested with both standard runs and environmental samples. Sample components were successfully identified by an independent computer search program with an accuracy approaching 80%.

This report was submitted in fulfillment of contract number 68-02-2745 by the Research Triangle Institute under the sponsorship of the U.S. Environmental Protection Agency. This report covers a period from September 2, 1977 to March 1, 1979, and work was completed as of March 1, 1979.

CONTENTS

Abstract iii
Figures vi
Tables viii
Charts ix
Abbreviations x
1. Introduction 1
2. Conclusions and Recommendations 3
3. Literature Search of Existing Deconvolution Methods 5
4. Implementation and Testing of CLEANUP. 24
5. Implementation of the CLEANUP Program on the UNIVAC Computer 67
6. User Description of Program Operation 76
7. Delivery to EPA 79
References 80

Appendices

A. CLEANUP Source Listing 81
B. "Extraction of Mass Spectra Free of Background and Neighboring Component Contribution from Gas Chromatography Mass Spectrometry Data" 151

FIGURES

Name | Page
1 Schematic representation of 14 scan window for masses 40-450 in core - a typical case 9

2 A typical singlet peak model for a given mass 10

3 Various types of situations which do not quality in histogram calculation 11

4 Peak detection histograms for sample processing window of Figure 1 12

5 Statistics of TIC vs scan number 13

6 Example of background subtraction for model peak 15

7 Peak model resolutions: two possible situations 18

8 Overall schematic of computer assisted qualitative analysis of GC/MS data 27

9 Five "raw" spectra used as input to CLEANUP 32

10 Resultant "cleaned-up" spectrum from Figure 9 33

11 Precleanup - A, postcleanup - B 34

12 TIC plot of brominated standards 36

13 Hardware TIC plot of GC/MS analysis of organics from coal pyrolysis 39

14 Typical precleanup of mass spectrum 40

15 Same spectrum after CLEANUP 40

16 Chemist's Printout from CLEANUP 44

17 Total ion current plot from test run 48

18 Graphic representation of possible settings of NTHIRD 60

19 Function of ITOM histogram window setting 61

20 Effects of linear error ratio 64

21 Effect of SATVAL on CLEANUP 66

22 Modules required for transfer of CLEANUP program to UNIVAC . . . 68

23 Raw TIC of INCOS test data 71

24 Possible consequences of scan rate too rapid for CLEANUP 72

25 Chemist's printout of EPA test run generated at TUCC 73

26 Chemist's printout of EPA test run 74

27 Example of CLEANUP control cards for EPA UNIVAC 77

TABLES

Number Page

1 Subroutines of CLEANUP by Function 20

2 Results of Comparison of Cleanup with Brominated Standards Searched for 26,209 Standards 37

3 Typical Printout for Identification of Coal Components (After CLEANUP) 41

4 Comparison of Selected Spectra from Coal Gasification Sample Searched Before and After CLEANUP 43

5 Description of Types of Samples Investigated to Test Out CLEANUP 47

6 Comparison of Manual US Computer Identification of GC-MS Run. . 49

7 CLEANUP Input Paramter Summary 58

CHARTS

Number		Page
1	Summary of Calling Structure and Subroutine Function	19
2	Steps Involved in Processing of CLEANUP Test Data (1) Preparation for and Interface to CLEANUP Program	28
3	Steps Involved in Processing of CLEANUP Test Data (2) Preparation for and Interface to Search Program	29
4	Steps Involved in Processing of CLEANUP Test Data (3) Generation of Versatec MS Plots from CLEANUP Files	31

ABBREVIATIONS

GC/MS - Gas Chromatography/Mass Spectrometry
TIC - Total Ion Current
A/D - Analog to Digital
MS - Mass Spectrometry
TUCC - Triangle University Computation Center
GC - Gas Chromatograph(y)
R - Correlation Coefficient
H - Hit
M - Miss
I/O - Input/Output

SECTION 1

INTRODUCTION

Gas Chromatography/Mass Spectrometry (GC/MS) is an analytical method which has the potential capability of allowing the total automated identification of a complex mixture of pollutants. Implementation of such a technique depends on the availability of two types of new technology. The first, the automatic acquisition and storage of fast scan low resolution mass spectrometry data, has been available for some time. The first GC/MS system involving the continuous recording of mass spectral data from a mass spectrometer operating in cyclic mode was first reported by Hites and Biemann in 1968.[1] At the present time these systems are commercially produced by many manufacturers.

The availability of computer programs for the analysis of this kind of data is at present still very limited. Although software development in this field constitutes a very active research area (roughly 100 papers are published yearly), very few integrated systems have yet been developed which are generally useful. This is in part due to the large variety of computer types, each with its own individual characteristics, but more importantly due to a general unavailability of broad based algorithms with an effective philosophy of approach. As a result, the vast majority of mass spectral runs are still being analyzed by hand.

The problem of the development of a generalized system for data analysis has two aspects. The first is how to separate mathematically the raw mixed spectra into well resolved components. The second is to assign a chemical structure to each resolved spectrum. Several approaches have been tried concerning the second problem, and one based on the pioneering work of Grotch[2] had been successfully implemented and realized in the form of a library search program now resident at the RTI Mass Spectrometry Facility prior to this project. The early results of this search program however, had not been totally satisfactory, usually for the reason that the input spectra were not sufficiently "clear", to allow unique identification. This problem was in large part alleviated by combining the search algorithm with an appropriate cleanup procedure. The implementation of each of these two stages and their effective combination turned out to enhance the capabilities of the entire system.

After a prototype system was developed which was shown to be effective, the CLEANUP programs which were developed were adapted to the EPA environment. Since manipulation of large data bases are much more easily accomplished in a large computer, it was felt that this process was best carried out by adapting data acquired using EPA's INCOS data system and processing them on a large computer. This approach had the dual advantage of making the programming effort much simpler, since core size limitations did complicate the programming, and in addition the processing could be accomplished independently, thus freeing the INCOS system to do other work. Again, the fact that programs were written in standard FORTRAN facilitated their transfer to the EPA UNIVAC computer.

Completion of the project by delivery of the program and its documentation to the contracting agency, is effected with this report and the accompanying listings (Appendix 1).

SECTION 2

CONCLUSIONS AND RECOMMENDATIONS

The work accomplished under this contract can be divided into four parts: (I) a literature search to determine the most effective algorithm for deconvolution of mass spectra; (II) implementing that algorithm on an existing computer and performing detailed tests of its efficiency, both from a chemical and machine standpoint; (III) making the program operational on the EPA UNIVAC computer installation using for input a tape created using the INCOS data acquisition system; (IV) making the necessary documentation and source code (Appendix 1) available to the contacting agency.

Three techniques were considered in Phase I: background subtraction, mass max analysis, and peak profile analysis. The peak profile analysis proposed in a recent article by Rindfleisch (Appendix 2) was ultimately chosen because of its demonstrated effectiveness in existing systems (the CLEANUP program) and because it incorporated the notions of background subtraction and mass max analysis as a starting point.

Once the Rindfleisch algorithm was chosen, the task of implementing that algorithm on the TUCC 370 was undertaken. First, a detailed study of the published algorithm was made in order to gain a better understanding of what was involved and to determine if improvements were necessary. Next, a copy of the Rindfleisch program modified to run on an IBM 370 was obtained through the cooperation of a group of chemists at the NASA Ames Air Force Base. This copy was then modified to run in a batch environment at TUCC and was subjected to extensive testing, both from a machine and chemical standpoint. This included running the program under the WATFIV compiler to insure program accuracy and machine independence, followed by testing with data acquired by a VARIAN CH-7 mass spectrometer. The test data consisted of three brominated standards, followed by fifteen runs involving a wide variety of environmental samples. The RTI library search program was used throughout to evaluate the output from CLEANUP. The laboratory standards were all correctly identified and correct identifications, approaching 80% were obtained on some environmental samples. In addition, some of the CLEANUP input parameters were studied in order to determine optimum settings.

Phase III, making CLEANUP operational in the EPA environment, consisted of three parts: obtaining a successful run of INCOS test data supplied by EPA on the TUCC 370, obtaining a successful run on the same data using a version of CLEANUP adapted to run on the EPA UNIVAC computer, and finally verifying that the output tape was compatible with the contracting agency's INCOS data acquisition software and yielded plausible output spectra.

This report, together with the accompanying listings and the installing of CLEANUP on the UNIVAC comprise the completion of Phase IV of this project.

It is now necessary to carry out a thorough testing of the programs to determine what conditions and sets of parameters must be used to obtain optimal results. The principles used in developing the software are sound, however, the programs need to be tuned to fit the acquisition conditions. Some attention should be given to control studies using known mixtures, because these give the best opportunity for evaluation of the system. A statistical evaluation of the results, especially when compared with other methods, would be very profitable.

SECTION 3

LITERATURE SEARCH OF EXISTING DECONVOLUTION METHODS

Phase I of this contract consisted of a literature search to consider available alternatives for the treatment of mass spectral mixtures, and thus arrive at a best method for deconvolution. Three general avanues of attack were examined: background subtraction, mass max, and peak profile analysis.

BACKGROUND SUBTRACTION

The technique of background subtraction is basically as follows. In its simplest form, a spectrum from either an unknown GC run or some blank run is chosen as a "background" spectrum. This spectrum is then subtracted from the unknowns within the run, and the resulting difference spectra are used for mass spectral identification.

In a slightly more refined version of the technique, the background spectrum is determined manually by visual inspection of a raw total ion current (TIC) plot, and in some instances several background spectra are chosen for use in different ranges of the chromatogram.

The difficulties associated with this approach are fairly obvious. Background spectra are not constant, either qualitatively or quantitatively. Thus, what constitutes the background in one section of the run may be very different from that in another. Additionally, background peaks arising from contaminants such as column bleed change in intensity especially during temperature programmed runs.

Thus the question of how the background subtraction is to be carried out is very difficult to answer. Even the mechanics of the subtraction can be done in a number of ways. Several unsophisticated schemes exist in which the background spectra are subtracted, peak for peak, from the raw digitized unknown. Invariably, some spectra are undercorrected, while in others, real data is eliminated from the output, making identification very difficult, either by hand or by machine. Human interpreters often prefer to look at raw data over background subtracted, since they can learn to ignore certain peaks without the removal of key unknown peaks by a clumsy program.

More sophisticated approaches have been tried wherein only a fraction of the background spectrum is subtracted. The problem of course, is to determine the proper size of the fraction.

All in all the background subtraction approach was eliminated since it is not sufficiently flexible, can give erroneous results, and lacks a way in which a consistent algorithm can be proposed for automatic processing.

MASS MAX ANALYSIS

A second and more sophisticated approach to mass spectral deconvolution has been taken by Biemann and co-workers.[3] This approach makes use of the principle that when a compound elutes from a GC column and is analyzed by continuous scan MS, the ions of each component show maximum intensity at some mean elution time.

The algorithm chosen, then, searches through each single ion trace for points of maximum intensity. The deconvoluted spectra are then defined as that collection of ions which have shown an intensity maximum within a given time window.

This approach was designed to allow the separation of mixtures of unknowns which are not sufficiently resolved to show individual TIC peaks, but which can be distinguished by the presence of single ion maxima which are not influenced by other ions not present in that particular unknown.

This idea is sound, and in fact, constitutes the first part of the Rindfleisch[4] procedure, described below. The procedure is very rapid and has actually been incorporated into other commerically available deconvolution software systems.

The program is deficient because it does not take into account a number of facts.

1. If a given ion shows a maximum at one point, it does not by any means guarantee that this ion will not be present in the other component.

2. If, as often occurs, an ion is present in both components in approximately equal amounts, the observed maximum may occur between the maxima of either component and be missed entirely by the system.
3. The technique does not take into account contributions from background noise or column bleed, and treats each ion in an all or nothing manner, and therefore cannot get a proper estimate of the intensity of peaks which derive from more than one source.

The method seems to have some merit for the purpose of identifying the presence of unresolved multiplets. However, it cannot accurately predict the actual spectra, and because of the deficiencies mentioned in the previous paragraphs, sometimes fails entirely, showing specious peaks, or missing multiplets.

PEAK PROFILE ANALYSES

The algorithm for the CLEANUP program herein described was originally due to Rindfleisch and co-workers at Stanford.[4] Their program was part of an integrated GC/MS data reduction system which ran on a PDP-11/20. Following is a detailed description of the elements and mode of operation of this program. This information was gathered from the published work on the subject as supplemented by our copy of the listings and other documentation.

Raw spectra output from GC/MS systems differ from the library spectra of pure components for at least three definable reasons. CLEANUP attempts to correct mathematically these spectra in three distinct ways, corresponding to these known error sources. These are 1) Background noise and column bleed, 2) Signal saturation, 3) overlapping peaks. We will call the three portions of the program dealing with these problems Tasks 1, 2, and 3.

Errors arise for the following reasons. Superimposed on the eluant spectrum there will always be a certain constant level of background noise, and in addition other extraneous signals which normally rise slowly with increasing column temperature due to the elution of some of the liquid phase of the column. Task 1 is concerned with the distinction between the origin of these kinds of signals. Task 2 tries to reconstruct the actual intensity level of peaks which are so strong that they have exceeded the range of the A/D convertor. This is the problem of peak saturation. Task 3 is the reconstruction of peak intensities of two or more compounds which elute so closely to one another that their mass spectral peaks are intermixed. The guiding principle which allows the CLEANUP program to function is that it uses information both in advance and after a particular scan to distinguish between real peak unknown information and extraneous data. Such issues as column bleed detection, doublet resolution, background detection, etc. are defined in the context of a region of scans rather than a single scan. The CLEANUP program makes use of fourteen MS scans in core at any one time (Figure 1). For tasks 1 and 2, only the seven innermost scans are used. For some details (e.g., background estimation), all 14 scans are needed. For multiplet resolution, all 14 scans are used. The processing proceeds from the beginning of the run to the end in a single pass always using 14 spectra. The spectrum window in core is progressively updated by the addition of one spectrum from one end and the dropping of one spectrum from the other.

The CLEANUP algorithm can be divided into stages from a programming point of view. The first stage is the examination of the spectral window and determine whether or not an eluant is present at all. When a peak, or group of peaks is detected, then tasks 1, 2, and 3 can be undertaken to improve the quality of the raw data. The detection process is accomplished as follows: The program looks at a window seven scans wide as the input spectra move through the fourteen spectrum window in core. For each of the 411 masses in the seven scan window (each scan covers masses 40 to 450 the program does a primitive check to see if the window contains a simple maximum (Figure 2). As maxima are detected, histogram of the frequency of "singlet" maxima is created. Data are ignored if there are any "lumps" (indicating a doublet) or if peak saturation is observed anywhere in the window (Figure 3). In addition to the histogram of the number of occurrences of maxima (Figure 4), a record of a slightly modified total ion current is kept (Figure 5), i.e, once a singlet maximum is detected, a rudimentary approximation to column bleed and background is subtracted (corresponding to the lowest value in the 14 a.m.u. window), the ion current at the maximum is estimated via interpolation, and the resulting ion current is added into a peak amplitude record. When an appropriate number of scans have passed through the window (the statistics are updated continuously) a decision is made on the basis of the number of maxima observed and the total amplitude of those maxima whether an eluant is present. It should be mentioned here that at this

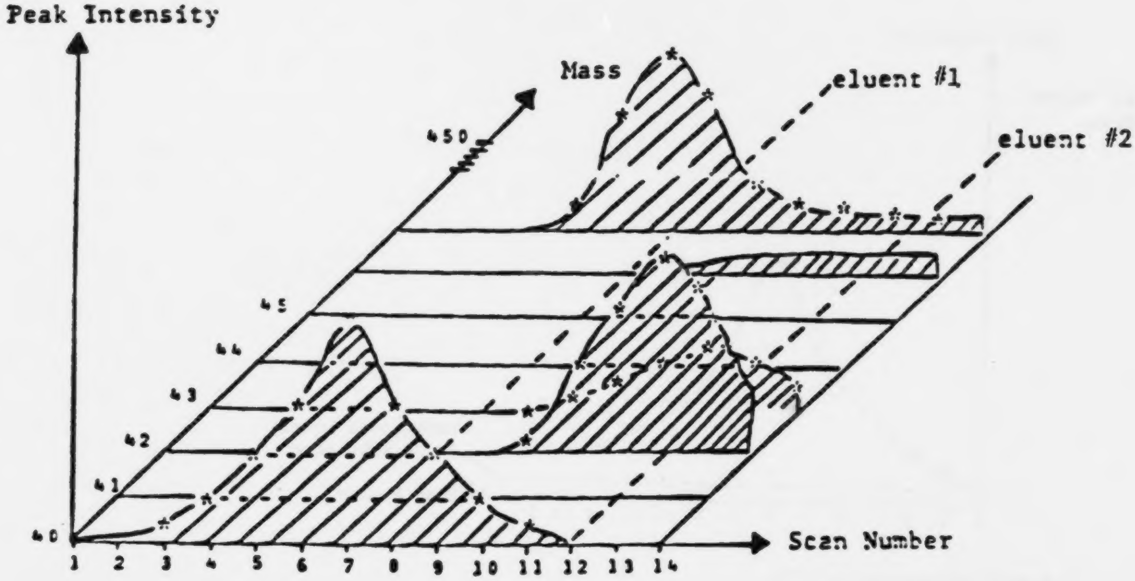

Peak Intensity
Mass
eluent #1
eluent #2
450
45
44
43
42
41
40
1
2
3
4
5
6
7
8
9
10
11
12
13
14
Scan Number

Figure 1. Schematic representation of 14 scan window for Masses 40-450 in core - a typical case.

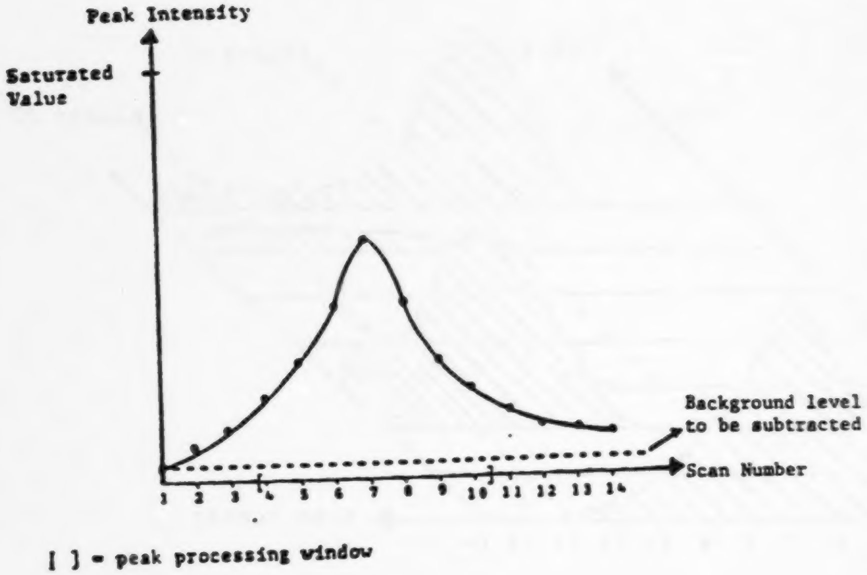
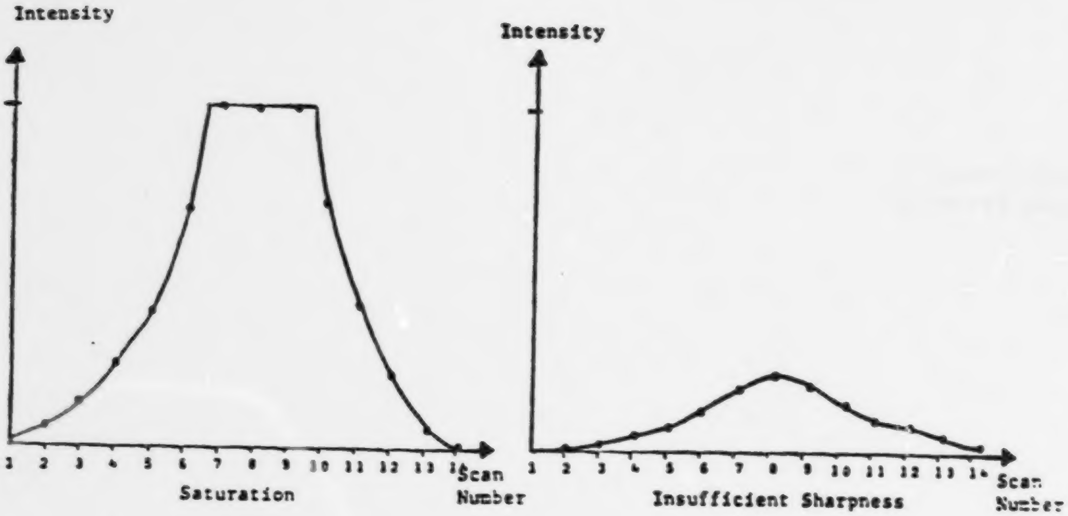
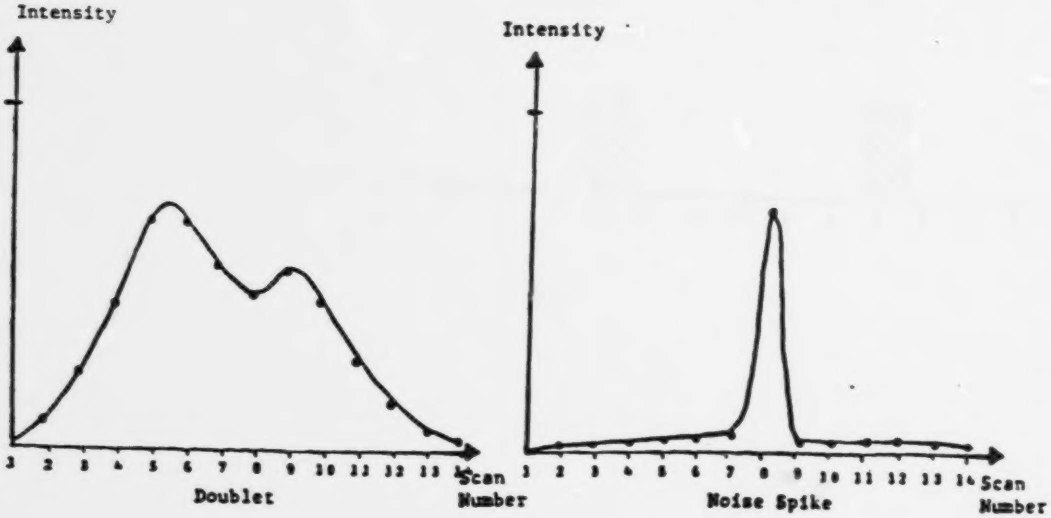


Figure 2. A typical singlet peak model for a given mass.



Saturation

Insufficient Sharpness



Doublet

Noise Spike

Figure 3. Various types of situations which do not qualify in histogram calculation.

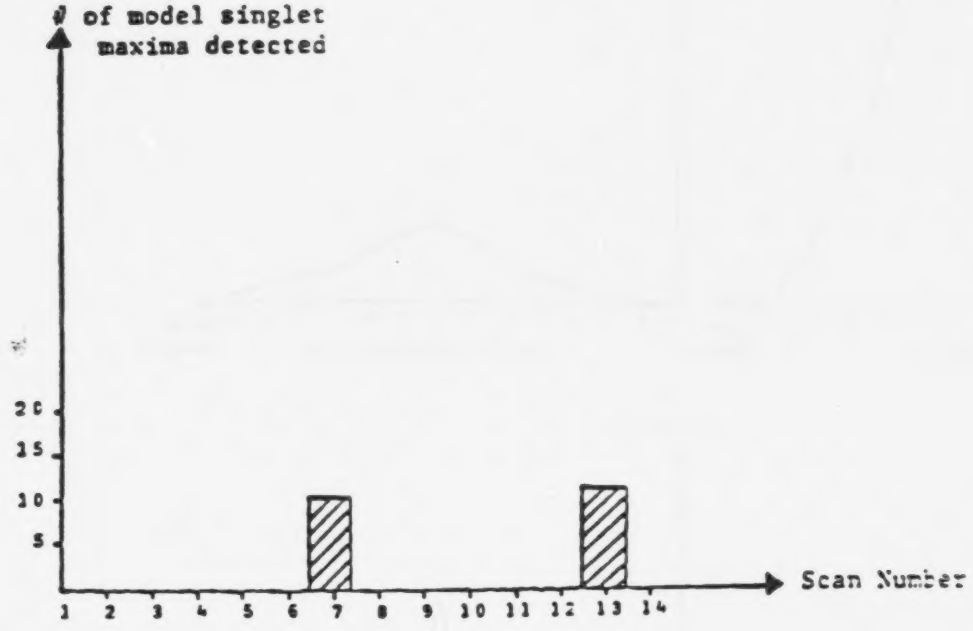


Figure 4. Peak detection histograms for sample processing window of Figure 1.

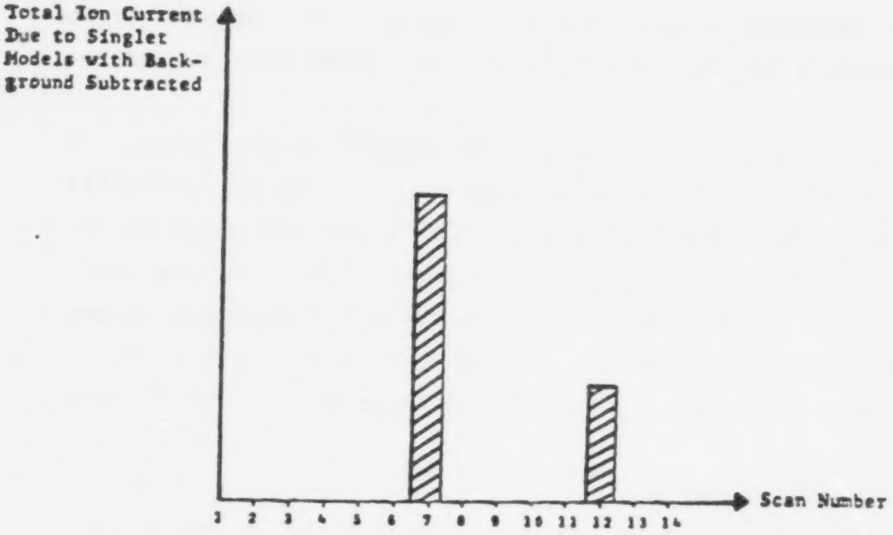


Figure 5. Statistics of TIC vs. scan number.

point in the program the "sharpness" of each candidate peak is computed. This number is a numerical approximation of the theoretical expression

$$\int \left| \frac{dy}{y} \right|$$

which is a measure of the relative steepness of each individual peak. It is to be noted that this number is independent of the height of the peak. There is a provision in the CLEANUP program for the rejection of peaks not meeting a minimum sharpness threshold (Figure 3). Noise spikes are eliminated by the property of the steepness function which favors wider peaks over very narrow ones.

Once the presence of a peak is detected, the CLEANUP program begins its real work, namely, the deconvolution of the peak mixtures and the accounting for bleed, noise, etc. The CLEANUP program assumes linear superposition of ion currents due to different causes, i.e., at any given point in time the total ion current can be represented as a simple sum of a background, column bleed, and whatever eluants are present. For the case of a singlet, the total ion current for a given mass Y(t) may be represented as a sum of three functions

$$Y(t) = P(t) + dt + c,$$

when the term dt represents the contribution to column bleed, c represents the contribution due to background noise, and P(t) is the peak contribution due to the signal. The problem lies in making an initial choice for the function P(t). The method used by CLEANUP is to examine the mass in the window with the sharpest non-saturated singlet maximum. This mass was stored by the previous phase of the program. The column bleed and background are then subtracted out (see Figure 6). Once the functional form of the model peak for a given mass is known, the contribution to the ion current due to the eluant compound at other masses can be computed via a least squares process, i.e., for any given mass

$$Y(t) \approx pP(t) + dt + c$$

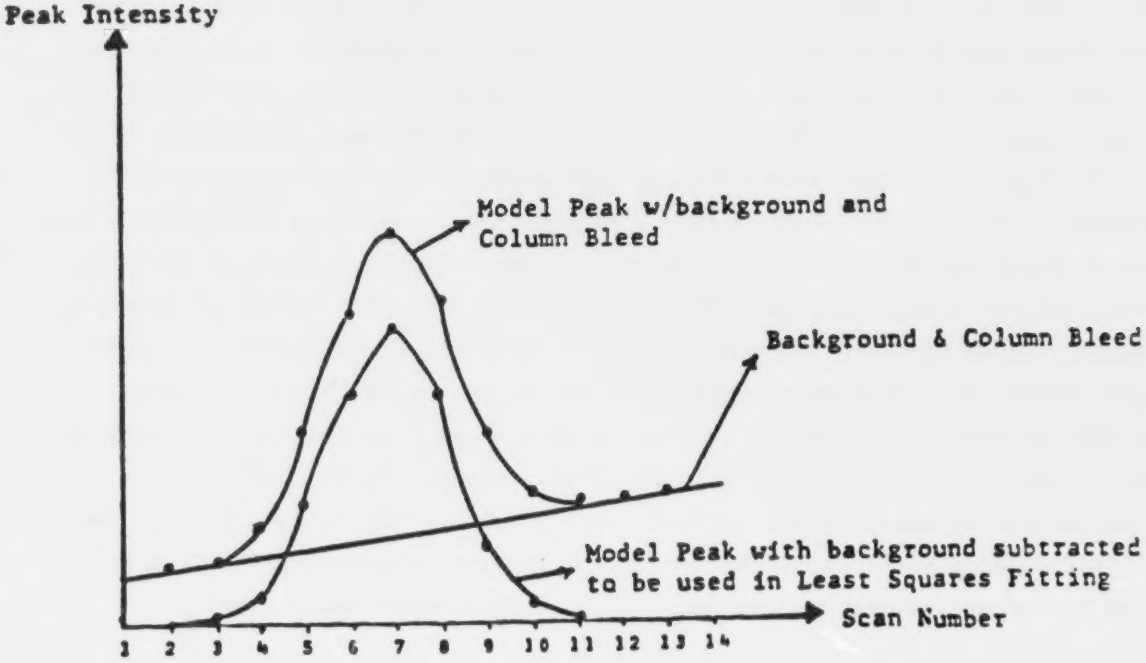


Figure 6. Example of background subtraction for model peak.

Here $Y(t)$ is the total signal for that mass, which is made up of a contribution of the signal, column bleed, and background. The aim is to minimize

$$D = \quad [Y(ti) - pP(ti) - dti - c]^2$$

in the interval covered by the peak, subject to the conditions

$$\frac{\delta D}{\delta p} = 0; \ \frac{\delta D}{\delta d} = 0; \ \frac{\delta D}{\delta c} = 0$$

This is done by a standard method.[5] It is important to realize that p, d and c are being varied in order to find the best possible match of three functions of fixed form (the model function P, a line through the origin, and a constant) to the given Function Y. The idea is that for a given eluant, all peaks will have the same functional form $P(t)$ and that peaks of different masses corresponding to that eluant differ only in relative amplitude - something which is taken account for in the factor p in $pP(t)$. Once the model function $P(t)$ is determined for a mass that is relatively free of background noise, neighboring eluents, column bleed, etc., a meaningful interpolation can be made for those masses which are contaminated with noise and bleed, or which are saturated. In the case of saturated peaks, a routine is used which interpolates an approximation of the functional form of $P(t)$ to obtain correct signal contributions for saturated points based on the non-saturated points on the fringe of each peak.

For doublet resolution, the model peak for the neighboring spectrum is added to the approximating function. $Y(t)$ is then approximately by

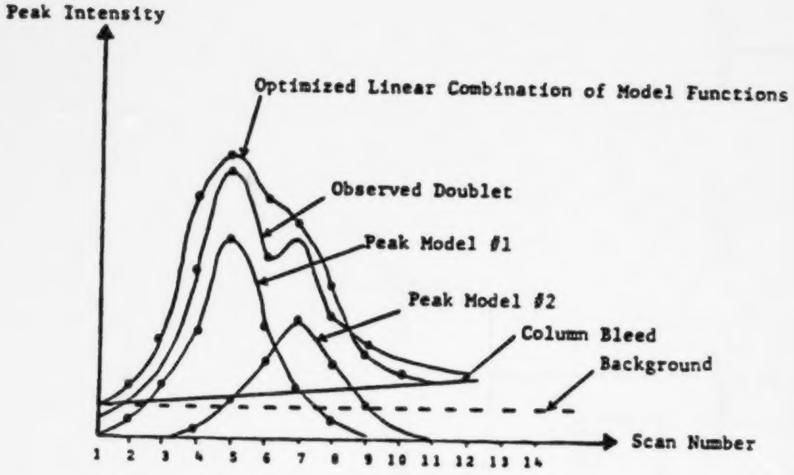
$$Y(t) \approx pP(t) + rR(t) + dt + c$$

where $R(t)$ is the model function for the second peak and r is its corresponding parameter. The idea is graphically illustrated for several cases in Figure 7.

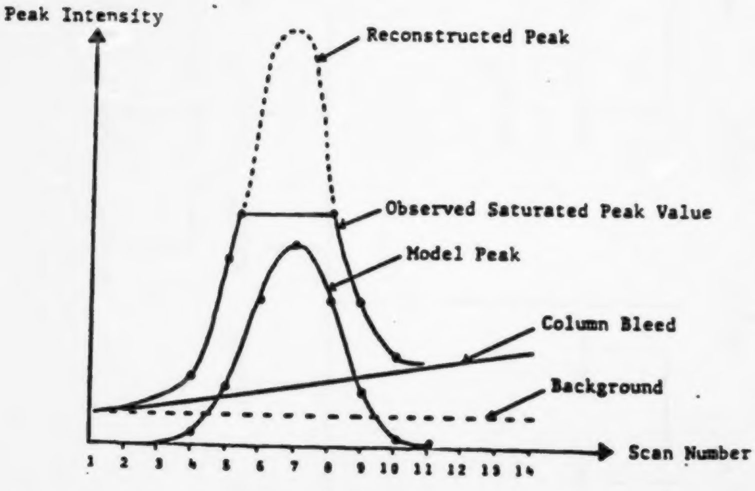
A diagram of the detailed subroutines of the CLEANUP program, along with a short description of these function is shown in Chart 1 and Table 1.

SCOPE AND LIMITATIONS

Some comments regarding the scope and limitations of the CLEANUP program are in order.

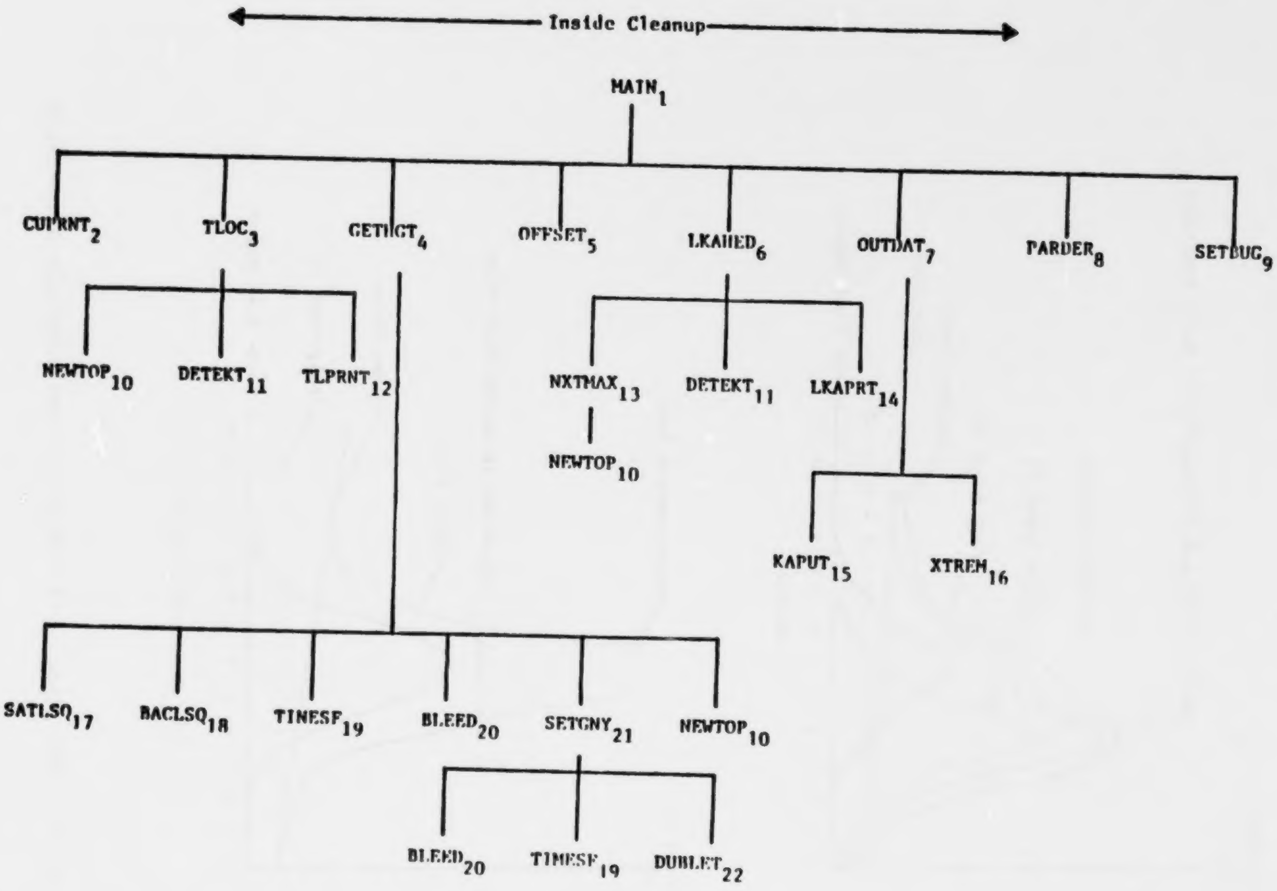

Peak Intensity
Optimized Linear Combination of Model Functions
Observed Doublet
Peak Model #1
Peak Model #2
Column Bleed
Background
Scan Number
1 2 3 4 5 6 7 8 9 10 11 12 13 14

Doublet


Peak Intensity
Reconstructed Peak
Observed Saturated Peak Value
Model Peak
Column Bleed
Background
Scan Number
1 2 3 4 5 6 7 8 9 10 11 12 13 14

Saturated Singlet

Figure 7. Peak model resolutions: two possible situations.



Inside Cleanup
MAIN$_1$
CURRNT$_2$
TLOC$_3$
GETHGT$_4$
OFFSET$_5$
LKAHED$_6$
OUTDAT$_7$
PARDER$_8$
SETBUG$_9$
NEWTOP$_{10}$
DETEKT$_{11}$
TLPRNT$_{12}$
NXTMAX$_{13}$
DETEKT$_{11}$
LKAPRT$_{14}$
NEWTOP$_{10}$
KAPUT$_{15}$
XTREM$_{16}$
SATLSQ$_{17}$
BACLSQ$_{18}$
TIMESF$_{19}$
BLEED$_{20}$
SETGNY$_{21}$
NEWTOP$_{10}$
BLEED$_{20}$
TIMESF$_{19}$
DUBLET$_{22}$

Chart 1. Summary of calling structure and subroutine function.

TABLE 1. SUBROUTINES OF CLEANUP BY FUNCTION

1. MAIN
   is the root segment which calls all other routines.

2. CUPRNT
   prints some debugging messages for MAIN.

3. TLOC
   builds histograms for eluant detection, and selects the model peak for each compound. A model peak is the singlet which is sharpest, and its mass fragmentogram is 9 spectra wide. It is important to note that histograms contain "singlet" information, and we use this information to detect eluants.

4. GETHGT
   is called after a spectrum has been detected, and is the executive for building the cleaned up spectrum from the raw data.

5. OFFSET
   is called only in the event that an overlapping situation arises, and aligns the model peaks with respect to their modes in arrays GMN and HNU in anticipation of processing a doublet.

6. LKAHED
   is called only if an eluant is detected. It looks ahead to see if another is present in the 14 spectrum window, and if so assigns the model peak and sets the flag OVLAP.

7. OUTDAT
   prints report information for the run and writes the clean up spectra to disk.

8. PARDIR
   is the operator communication subroutine. It sets some parameters which can be entered either from cards or from a terminal.

9. SETBUG
   does some initialization of output buffers, and reads in the first 13 spectra.

10. NEWTOP
    finds the actual mode and the displacement of this mode from the assumed mode as well as computing the integral $|dy|/y$ (rate) from a singlet peak profile.

(continued)

TABLE 1 (continued)

11.	DETEKT	examines histograms to determine if an eluant is present. The major criterion for an eluant to be detected is that its histogrammed total ion current be greater than MINTIC.
12.	TLPRNT	prints some debugging messages for TLOC.
13.	NXTMAX	builds histograms in the look ahead mode. Due to core limitations, it can only look ahead 3 spectra, and in the extreme case, i.e., the 3rd spectrum ahead is determined to be an eluant, the determination is done with less information than its predecessors. If 16 spectra are in core, NXTMAX can easily be set up to use exactly the same criteria as TLOC in building histograms. Some heuristics have been put in TLOC and DETEKT to take care of the minimal cases. This minimal situation occurs when a very weak eluant is found ahead, and then passed up when it becomes a primary candidate.
14.	LKAPRT	prints some debugging message for LKAHED.
15.	KAPUT	outputs to the line printer a short summary of the results of the programs processing.
16.	XTREM	determines the minimum value and its subscript given an integer array and its length.
17.	SATLSQ	processes saturated peaks, and tries to compute actual peak heights by taking the ratios of nonsaturated amplitudes in the saturated peak profile with their corresponding amplitudes in the model peak.
18.	BACLSQ	does a least squares fit using the peak profile of a singlet, and a model peak to determine the singlets contribution to a detected spectrum.
19.	TIMESF	uses a parabolic interpolation to place a peak profile on the same time co-ordinate as the model peak for a detected elutant.
20.	BLEED	fits a line to a peak profile and uses goodness of fit, and average value of Y, where Y=mx+b, to see if the profile is column bleed.

(continued)

TABLE 1 (continued)

21. SETGHY

    is called only in the event of a doublet situation, and decides whether or not a given peak profile will be passed through the doublet resolver.

22. DUBLET

    is the doublet resolver, and does the least square fit of two model peaks, and a doublet peak profile and puts the second peak contribution in a different array.

The construction of model peaks for an eluant depends on at least one mass in the processing window having a well defined, well resolved singlet. This requirement will not be satisfied if two eluants are within roughly one scan time of each other, or if there is an unsaturated peak for use as a model, i.e. in extremely bad run conditions.

Also the CLEANUP procedure breaks down in the case of interacting elutents because the ion currents are not longer a simple sum. This is an inherent limitation of the linear reconstruction model.

Finally, the program currently lacks the capability of providing the user with a visual TIC plot. This is essential in order to allow the user to analyze the final results.

SECTION 4

IMPLEMENTATION AND TESTING OF CLEANUP

Phase II of the contract consisted of implementing a version of the Rindfleisch program on the TUCC 370 and testing its performance. This involved obtaining an existing copy of CLEANUP and modifying it to run at TUCC, developing support software to facilitate the analysis of CLEANUP output, and extensive testing of CLEANUP on a variety of environmental samples.

INITIAL PROGRAM MANIPULATION

The first step in implementation of the CLEANUP program was to obtain a listing and initial documentation from the author. This was achieved with relatively little difficulty. Dr. Rindfleisch and his colleagues were very generous in providing us with full listings and some documentation of their program. Since these programs were designed to run on a PDP-11 computer, they were written in a combination of PDP-11 assembler language and a PDP-11 version of Fortran, which would have made their translation into standard Fortran a very difficult task. We were informed by the Stanford group, however, that another group of chemists at the NASA Ames Air Force Base had been working on a Fortran implementation of the Rindfleisch program and we were referred to that group.

The Ames scientists were cooperative in forwarding their versions of the Rindfleisch programs to us, however it was immediately clear when we received tapes from them that there were large discrepancies between their programs and a working system. There were numerous errors in the program, indicating that they had not been fully tested.

After some additional conversations with the Ames group, an improved version of their Rindfleisch program was sent to us which we began to use. The new program had been chnaged, and some of the errors, detected previously had been corrected. This new program was then reorganized to make it operational in the TUCC environment. The mode of inserting variable data and options into the program was completely altered. The original programs were designed to operate interactively from a terminal, and the variables for each run were introduced in a conversational mode. If the results are not satisfactory, then the parameters are changed and the program re-run, etc. This process is continued until the operator is satisfied with the results, or feels that no further improvements can be made. Since in our environment programs must be run in a batch process, it was necessary to proceed somewhat differently. A set of trial parameters were obtained from the program authors, which were entered into the RTI program from cards. The effect of the program paramters on the output is discussed in 3.4.

The main body of the program was debugged by the use of the WATFIV compiler. This compiler is useful, not only because of its very well documented error messages, but because the compiler, by its nature, cannot operate using special system subroutines. This means that the entire program, if run under WATFIV, is insured to be compatible with other computers running with standard Fortran packages. By this procedure, the entire program was debugged and freed from a number of errors, particularly those relating to the initialization of variables prior to their use, which in our experience, has proven very difficult to trace down at later stages of program development.

SUPPORT SOFTWARE FOR CLEANUP

After the program ran freely, it was necessary to test it on some typical environmental data. In order to do this, GC/MS data which previously had been generated from environmental samples in the RTI laboratory were adapted to the CLEANUP program. This required the writing of several sets of programs: one translated raw mass spectral data into the input format required by CLEANUP; a second translated the output data from CLEANUP into a format required by the search program; a third set generated plot output from CLEANUP on an in-house VERSATEC plotter. A more detailed description of this software follows.

RTI Search System Interfacing

These programs accepted mass spectral data, which had been acquired on a Varian CH-7 mass spectrometer, and automatically selected spectra for introduction into a library search program. Selection of specific spectra was carried out by the analysis of the total ion current (TIC) profile of the GC run. In the case of spectra acquired using the Varian system, the TIC profile was very easy to analyze numerically, because the data system was designed to acquire TIC information simultaneously with MS information throughout the course of the run. By means of a multiplexer, the data system acquired ca. 100 TIC points, which were received from an ion current electrode placed at the exit of the source of the spectrometer, during each MS scan. These TIC points were recorded on the output tape in a Varian coded format and it was possible to regenerate this TIC information in both numeric and graphic form from the original tape.

Whereas most other GC/MS computer systems generate TIC data by the summation of ion intensities for each mass spectrum run (software TIC), the hardware TIC data is much richer in detail since it produces ca. 100 points per GC scan rather than one in the software method. This advantage is particularly important when data are being scanned relatively slowly as compared with the TIC peak width. The TIC trace generated from the data encoded on the tape was very useful to evaluate the CLEANUP program, since CLEANUP does not produce any visual record of the GC run, but simply selects specific peaks for subsequent evaluation, in our case by means of library search. We compared the selection of spectra made by analysis of the hardware TIC trace with those spectra chosen by CLEANUP. This approach was extremely useful in the evaluation of the CLEANUP program. Figure 8 shows a schematic diagram of the RTI data processing system.

Splicing the CLEANUP program into te RTI data processing system consisted of two parts: formatting the data generated by the MS-7 GC/MS system so that it could be read as input by CLEANUP (Chart 2) and formatting the output generated by CLEANUP so that it could be used as input to the RTI library search program (Chart 3).

As these programs were being written, a tape was prepared of a standard GC/MS run on the RTI MS-7 GC/MS system and was hand carried to TUCC. After a few trials, we were successful in creating CLEANUP-compatible input data, getting CLEANUP to run to completion on that data, and finally obtaining a successful run on that data with the RTI search program.

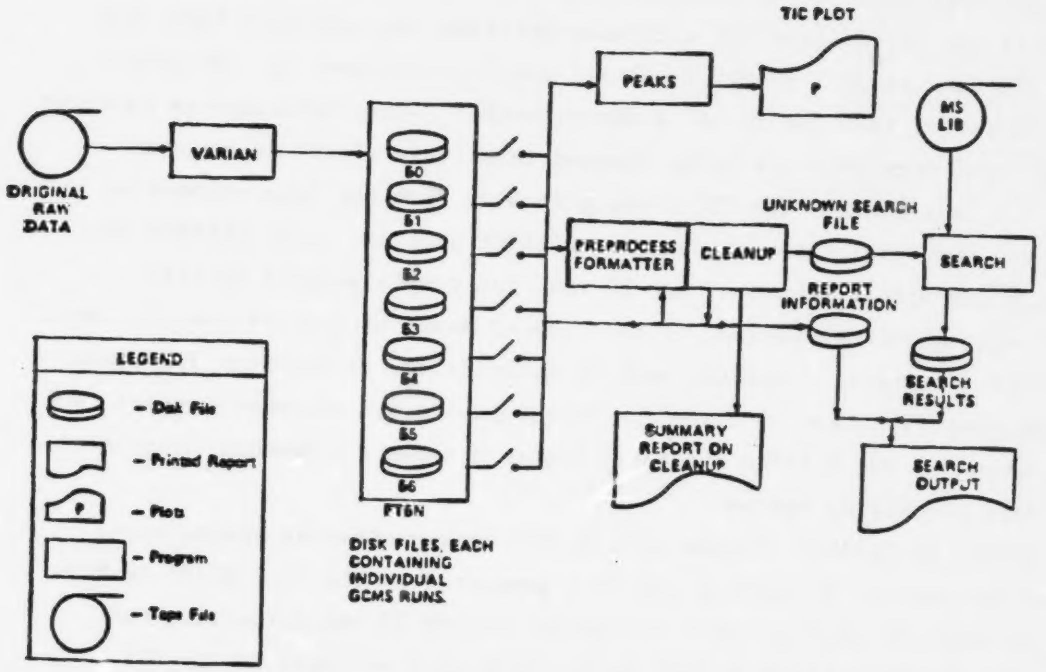
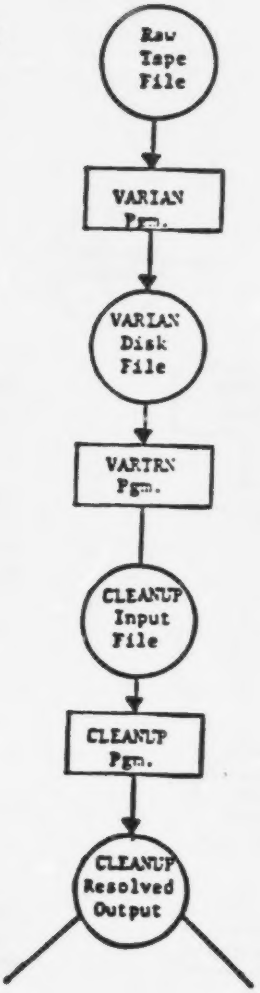

TIC PLOT
PEAKS
P
MS LIB
ORIGINAL RAW DATA
VARIAN
60
61
62
63
64
65
66
FT6N
DISK FILES, EACH CONTAINING INDIVIDUAL GCMS RUNS
UNKNOWN SEARCH FILE
PREPROCESS FORMATTER
CLEANUP
SEARCH
REPORT INFORMATION
SEARCH RESULTS
SUMMARY REPORT ON CLEANUP
SEARCH OUTPUT
LEGEND
– Disk File
– Printed Report
– Plots
– Program
– Tape File

Figure 8. Overall schematic of computer assisted qualitative analysis of GC/MS data.


Raw Tape File
VARIAN Pgm.
VARIAN Disk File
VARTRN Pgm.
CLEANUP Input File
CLEANUP Pgm.
CLEANUP Resolved Output

Files Generated

Varian data tape containing raw output from GC/MS

Intermediate step, output file from VARIAN

Output file of VARTRN in CLEANUP compatable input format

Output file of CLEANUP containing resolved spectra

Chart 4:
Display of input and output spectra on VERSATEC plotter

Programs Involved

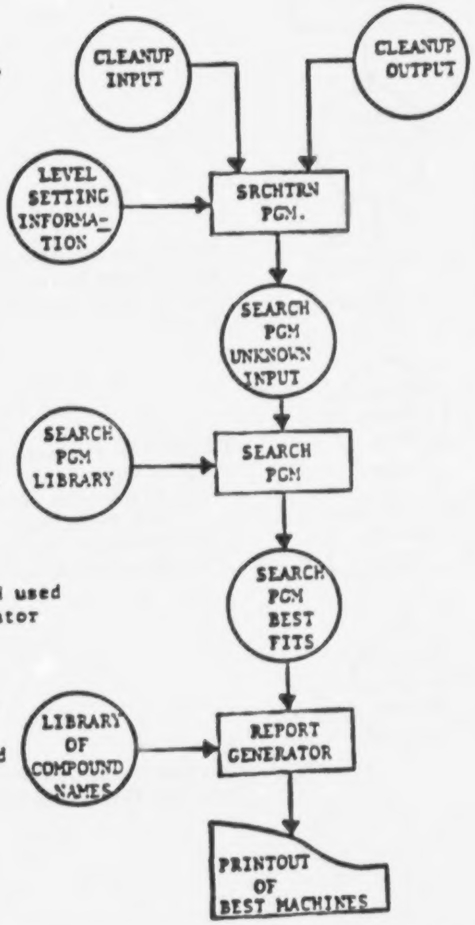
Translates raw Varian data to Fortran compatible format

VARTRN is an intermediate program which transforms the output of VARTRN into CLEANUP compatable input

CLEANUP is the main processing program which:
(1) Eliminates column bleed and background
(2) Reconstructs saturated peaks
(3) Resolves mixed doublets

Chart 3:
Examination of results when fed into RTI Search System

Chart 2. Steps involved in processing of CLEANUP test data (1). Preparation for and interface to CLEANUP program.


CLEANUP INPUT
CLEANUP OUTPUT
LEVEL SETTING INFORMA-TION
SRCHTRN PGM.
SEARCH PGM UNKNOWN INPUT
SEARCH PGM LIBRARY
SEARCH PGM
SEARCH PGM BEST FITS
LIBRARY OF COMPOUND NAMES
REPORT GENERATOR
PRINTOUT OF BEST MACHINES

Input and Output files for CLEANUP (chart 1)

Sets in tens levels of unknowns into appropriate octill bracket

SRCHTRN converts the output spectra from CLEANUP into SEARCH compatible format

SEARCH searches through a library (for each unknown) and saves the best fits

Search library containing some 26,000 compounds

Output of SEARCH used by report generator

File of library compound names which are looked up to generate report

REPORT generates a "chemists printout" of the results from search

"Chemists Printout" of best choices for unknowns

Chart 3. Steps involved in processing of CLEANUP test data (2). Preparation for and interface to search program.

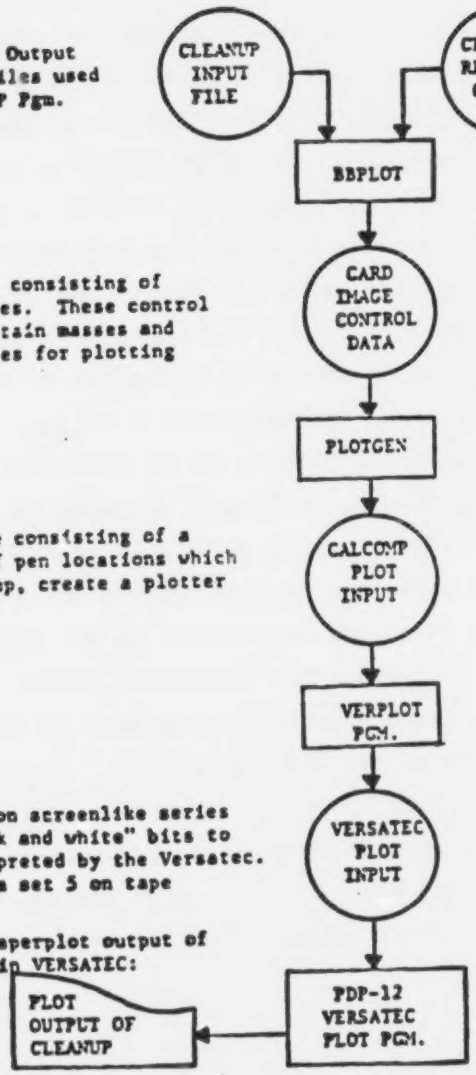
Plot Out for CLEANUP

At this stage it was necessary to carry out a critical evaluation of the results to determine the efficacy of the CLEANUP algorithm. Although the CLEANUP program executed successfully, preliminary examination of the printed data for a few isolated cases indicated that the initial parameters chosen were not ideal, and that although the program seemed to recognize the presence of individual components in the GC run, the output spectra were not yet suitable for characterization of the unknown GC component. In order to proceed in an orderly manner, then, it was felt that a method for displaying the results of CLEANUP needed to be developed. Software to generate plots from MS data have been available for some time, however, these programs generate output intended for the university operated CALCOMP plotters which were entirely too slow to generate the hundreds of plots needed in the course of this investigation. An alternative approach of writing software to generate plots on an in-house VERSATEC electrostatic plotter (several orders of magnitude faster than the CALCOMP incremental plotters) was adopted. The process consisted of two steps. First a program at TUCC read the CLEANUP output spectra as input and generated picture records which were then written to tape. The tape was then hand carried to the RTI laboratory and read by a PDP-12 program which generated the actual plots. This method proved very economical, since plots were generated at low costs (50¢ per picture) and high speeds (ca. 20/30 sec/plot). In this fashion it was possible to generate up to 100 plots in an hour, a task which would have tied up the TUCC network CALCOMP plotters for several days. The entire scheme for generating plots together with already existing software is shown in Chart 4 with sample output in Figures 9-11.

TEST RUNS WITH CLEANUP

Brominated Standards

A preliminary evaluation of the CLEANUP program using two GC/MS runs was made. The first was a relatively simple mixture of a series of


CLEANUP INPUT FILE
CLEANUP RESOLVED OUTPUT
BBPLOT
CARD IMAGE CONTROL DATA
PLOTGEN
CALCOMP PLOT INPUT
VERPLOT PGM.
VERSATEC PLOT INPUT
PDP-12 VERSATEC PLOT PGM.
PLOT OUTPUT OF CLEANUP

Input and Output spectra files used by CLEANUP Pgm.

Comparison spectra after completion of the CLEANUP Procedure

BBPLOT is an intermediate program which generates card image input for a plot generation program

Disk file consisting of card images. These control cards contain masses and intensities for plotting

Existing program to produce CALCOMP plots

Disk file consisting of a series of pen locations which as a group, create a plotter picture

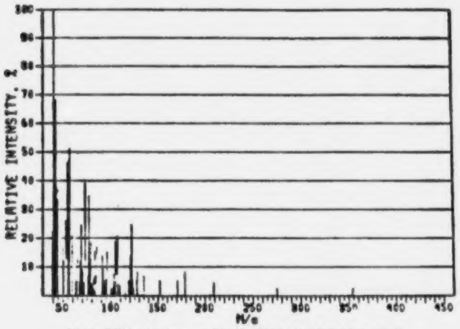
VERPLOT is a post processor for the TUCC plotter routines which translates CALCOMP compatable input plot records into VERSATEC compatable plot records

Television screenlike series of "black and white" bits to be intrepreted by the Versatec. This data set 5 on tape

Actual paperplot output of spectra in VERSATEC:

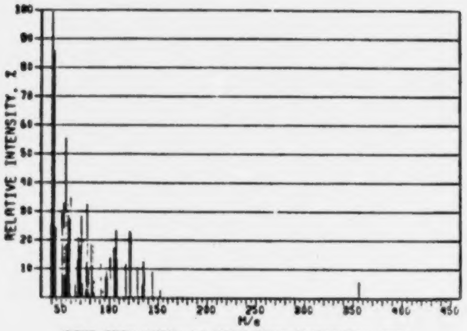
This program reads the VERSATEC plotter input on the inhouse tape drive and generates the actual plots on the VERSATEC.

Chart 4. Steps involved in processing of CLEANUP test data (3). Generation of Versatec MS plots from CLEANUP files.



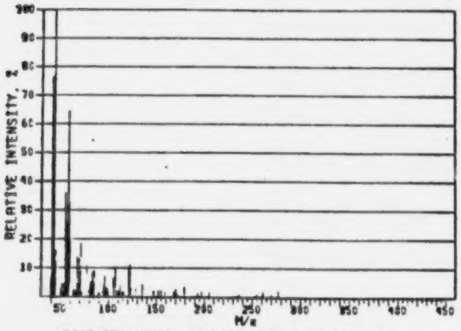
BASE PEAK MASS= 41,BASE PEAK INTENSITY= 265

NUMBER= 23302
RUN FOR WM F HARGROVE
03/23/78



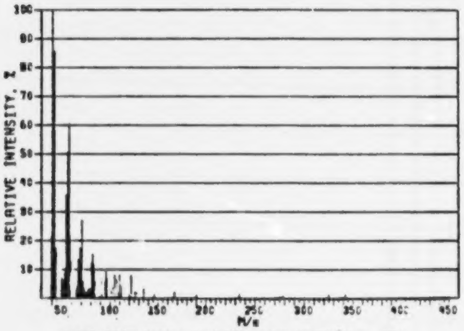
BASE PEAK MASS= 41,BASE PEAK INTENSITY= 256

NUMBER= 23303
RUN FOR WM F HARGROVE
03/23/78



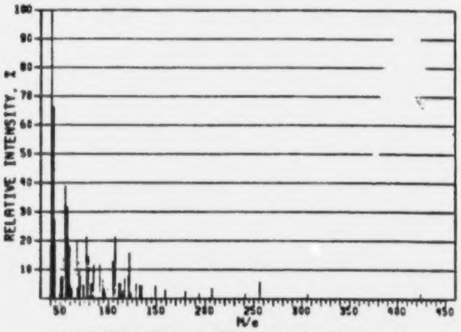
BASE PEAK MASS= 43,BASE PEAK INTENSITY= 665

NUMBER= 23304
RUN FOR WM F HARGROVE
03/23/78



BASE PEAK MASS= 41,BASE PEAK INTENSITY= 576

NUMBER= 23305
RUN FOR WM F HARGROVE
03/23/78



BASE PEAK MASS= 41,BASE PEAK INTENSITY= 411

NUMBER= 23306
RUN FOR WM F HARGROVE
03/23/78

Figure 9. Five "raw" spectra used as input to CLEANUP.

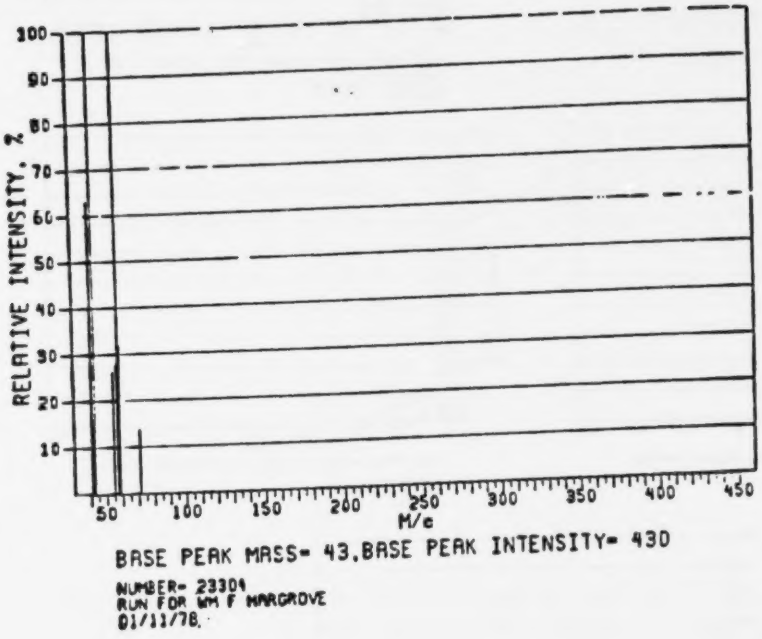


Figure 10. Resultant "cleaned-up" spectrum from Figure 9.

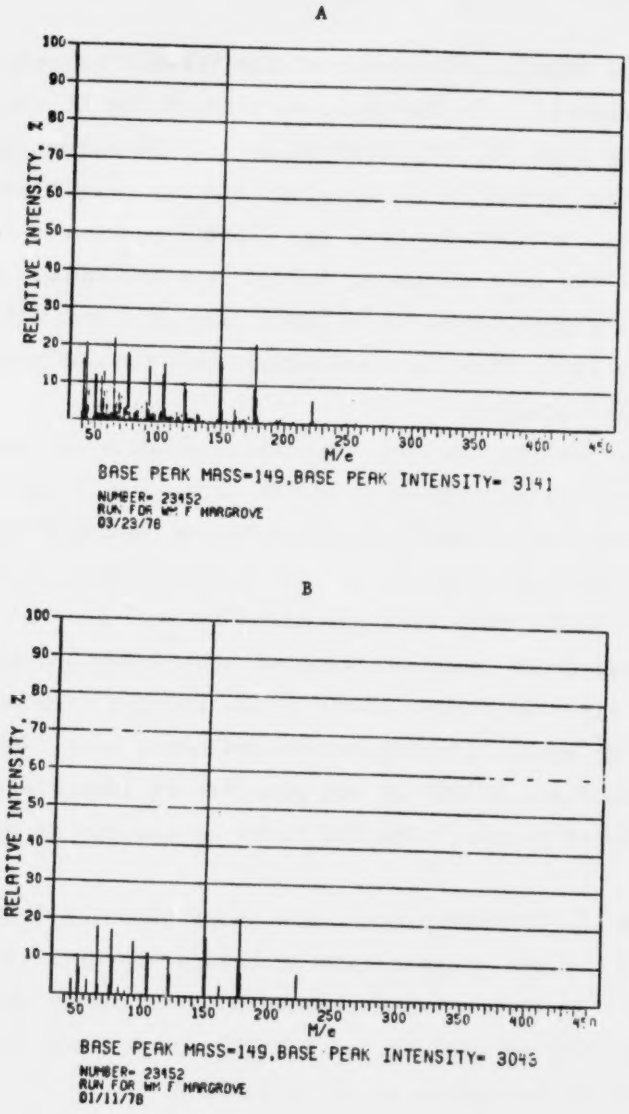


Figure 11. Precleanup - A, postcleanup - B.

brominated standard compounds. This series was chosen to serve as a control. The compounds in the run were known, and they were reasonably easy to identify.

In the case of the brominated standards, the CLEANUP procedure worked essentially perfectly. In the first section of the GC run the spectra were relatively free of background noise and the output spectra from the CLEANUP program resembled the original spectra very closely. This fact is not trivial in that it shows the CLEANUP procedure is essentially benign for the case of uncomplicated spectral data. Thus, if the original data are good, then the CLEANUP step does not alter them or degrade them in any way. This is a necessary condition of performance for a good CLEANUP program.

In the latter portions of the run, increasing amounts of background noise from column bleed became evident. In the case of weak spectra near the end of the run, the CLEANUP procedure became indispensible for the identification of several components. As it turned out, in addition to the brominated standards which were specifically put into the mixture, several additional components, many of them not containing bromine, were found. For example, spectrum number 23469 in the run had been seen many times previously during manual identification and never been identified. With the help of CLEANUP and SEARCH it was possible to identify this component as an alkylated phenol. The TIC trace of the run is given in Figure 12.

The results from the CLEANUP program are shown graphically in Figures 9 through 11. Figure 9 shows five successive spectra (the computer actually uses 7) from which a cleaned up spectrum is derived (Figure 10). Another pair of pre and postCLEANUP spectra are shown in Figure 11. A summary of the run is shown in Table 2.

Coal Run

The second GC/MS data set which was examined consisted of a low boiling fraction from a coal gasification process. With this data the situation was virtually the opposite from the previous case, *i.e.*,

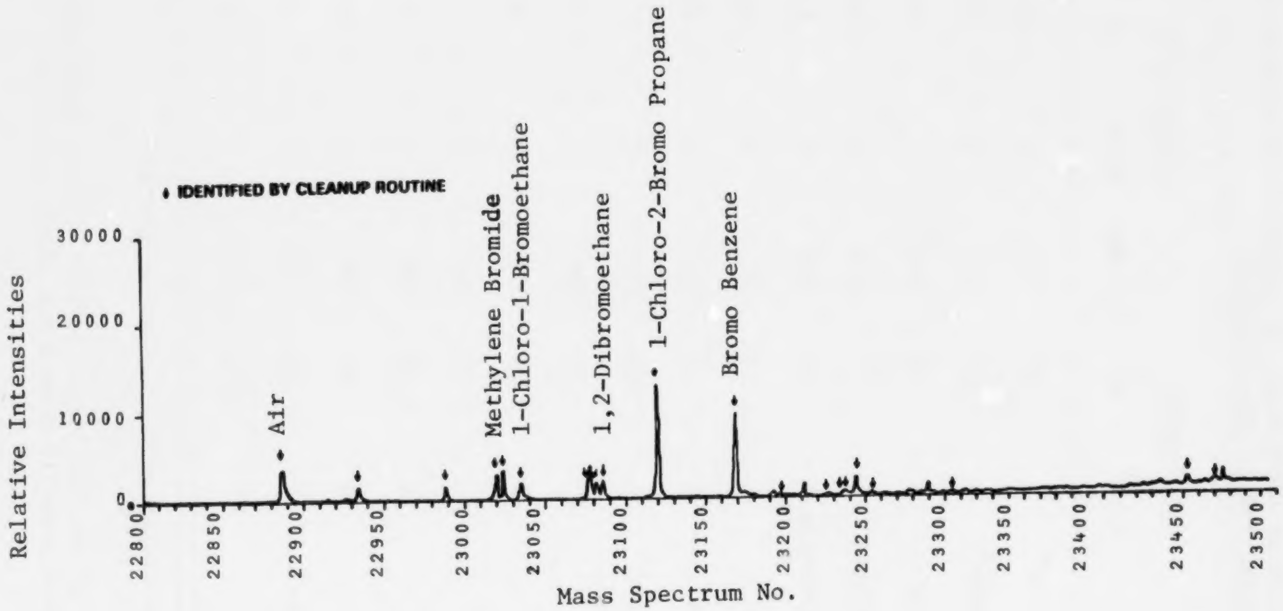


Figure 12. TIC plot of brominated standards.

Table 2. RESULTS OF COMPARISON OF CLEANUP WITH BROMINATED STANDARDS SEARCHED FOR 26,209 STANDARDS

		Precleanup			Postcleanup		Elution
	Compound	Rank[b]	R	TIC Max.	Rank	R	Time
1.	methylene chloride[a]	1	.31	752	1	.67	15.63
2.	1,2-dichloroethane[a]	1	.33	672	1	.54	21.93
3.	methylene bromide	1	.36	720	1	.49	25.55
4.	bromodichloromethane	1	.37	688	1	.43	26.00
5.	1-chloro-2-bromoethane	1	.30	704	1	.43	27.30
6.	_n_-butyl bromide	>5	<.19	512	2	.33	32.43
7.	ethylene bromide	3	.17	800	1	.31	33.25
8.	chlorobromopropane[a]	1	.28	1392	1	.37	37.22
9.	bromobenzene	>5	<.18	1088	2	.27	42.93
10.	dimethylphenol[a]	1	.31	320	1	.45	46.20
11.	_n_-butane[a]	>5	<.19	304	1	.53	49.35
12.	acetophenone[a]	>5	<.22	320	1	.39	51.10
13.	toluene[a]	>5	<.21	304	1	.60	52.97
14.	diethyl phthalate[a]	1	.33	478	1	.52	76.07
15.	octyl phenol[a]	>5	<.15	432	1	.52	78.05

[a] Compound was not knowingly part of standard mixture.

[b] The ordinal number of the similarity of the library component to the unknown. The program normally prints out the 5 best fits.

this sample consisted of a deliberately chosen, highly complex mixture, which would test the system to its limit to determine the most that could be expected from this program.

In this GC/MS data set, the chromatography conditions and the relatively slow sampling rate assured that the majority of MS scans would consist of mixtures, contain also background noise, column bleed, and other components in the system. Although the specific components in the mixture were not known, the nature of the sample gave strong clues as to what kinds of compounds were to be expected. A TIC plot of this run is shown in Figure 13.

Not surprisingly then, when raw coal run MS data were introduced into the MS search routines, the output was nearly totally useless, since the correlation coefficients were almost uniformly low and those compounds which were chosen as the closest fitting were obviously completely erroneous. We believe that we understand why this occurred. When samples are highly impure, numerous additional small peaks, particularly in the high molecular weight regions usually appear. The search system normally has no way of distinguishing these weak noise peaks from similar weak peaks which may be present in high molecular weight components. As a result, these peaks are picked up as possible molecular ions for high molecular weight components, and the results are entirely erroneous.

After the CLEANUP program had been applied to this set of data, the results, although far from perfect, improved noticeably. A sample of the search program output is shown in Table 3. Most of the high molecular weight noise was removed by CLEANUP, resulting in the identified compounds being of a much more "reasonable" nature. In addition, the search program's correlation coefficients rose dramatically, and indeed, most of the identified components seemed to be reasonable, based on the known origin of the sample. Thus hydrocarbons, low molecular weight phenols, and specific compounds containing heteroatoms such as thiophene and carbon disulfide were identified correctly by the search system with good correlation coefficients. We feel that this represents a remarkable achievement for such a difficult mixture. Figures 14 and 15 show a

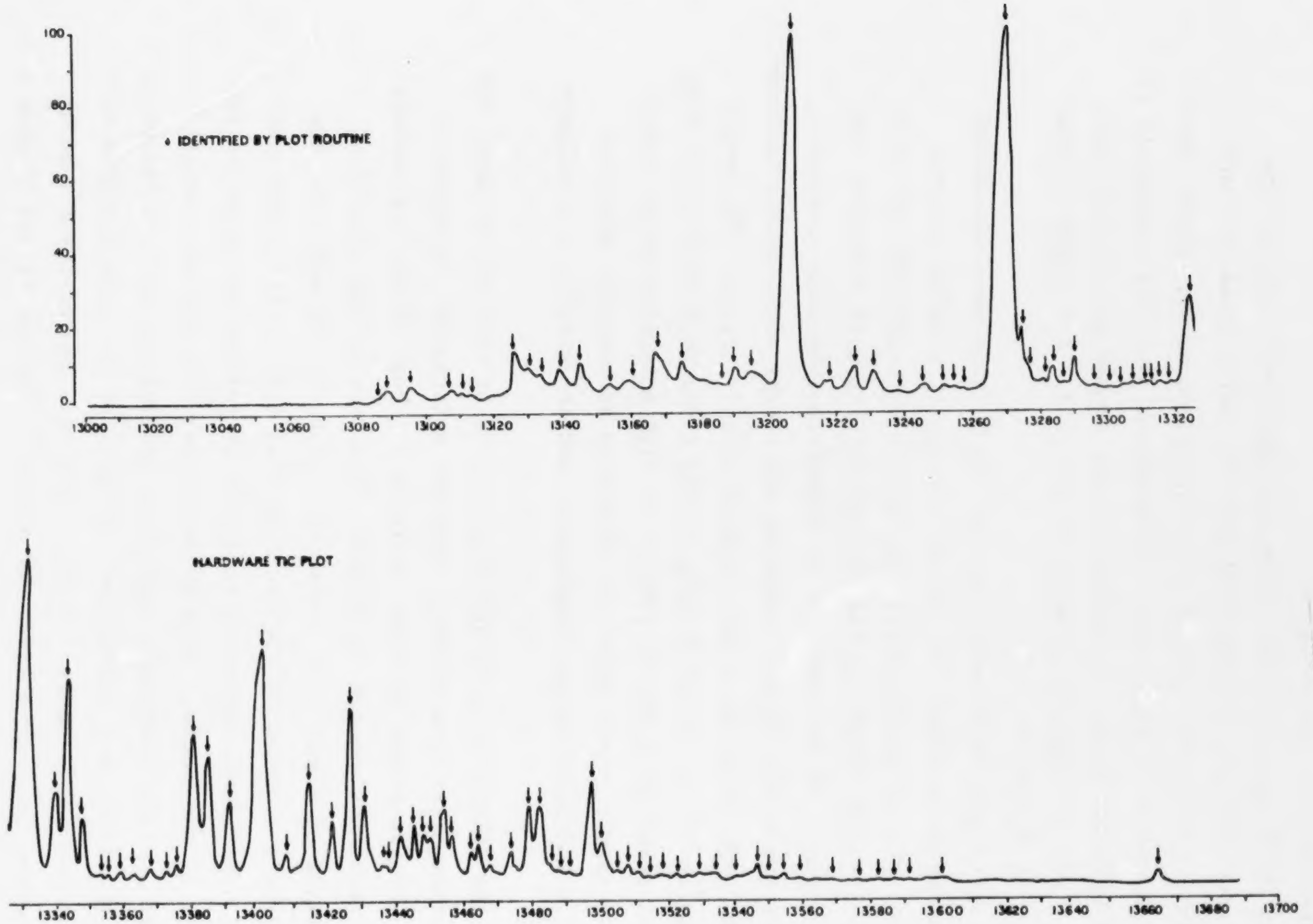


Figure 13. Hardware TIC plot of GC/MS analysis of organics from coal pyrolysis.

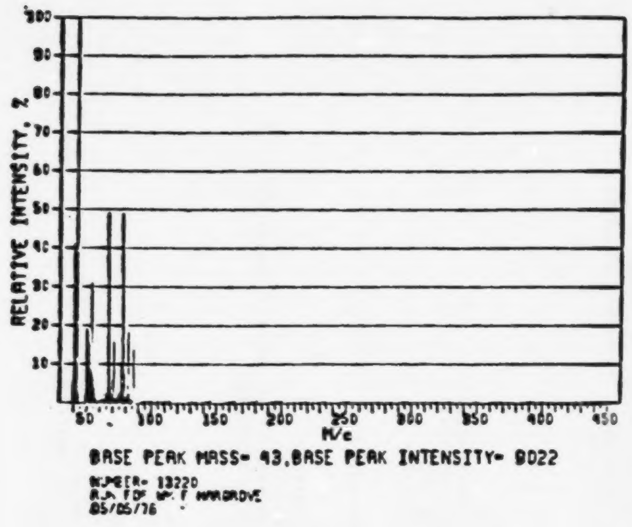


Figure 14. Typical precleanup of mass spectrum.

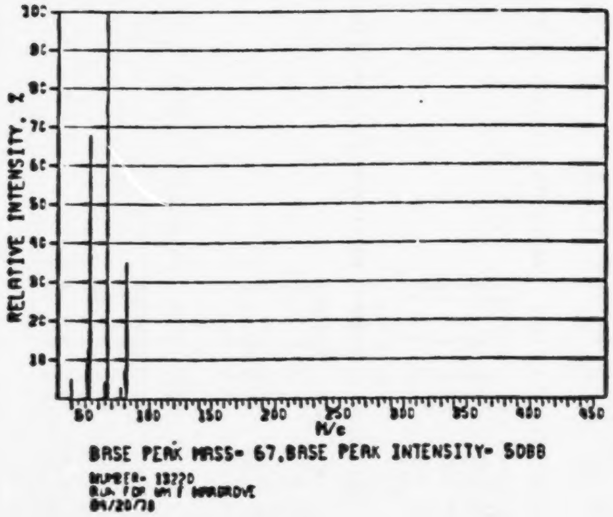


Figure 15. Same spectrum after CLEANUP.

```
SPEC # 4211  ELUTION TIME=   24.62 MIN  ELUTION TEMP=   0  TIC MAX=   928

     ID#    R      MWT    FORMULA                                NAME
1.    622 0.47     88 C   4 H  8 O  2                            ETHYL ACETATE
2.   2234 0.38    118 C   6 H 14 O  2                            1-ISOPROPOXY-2-PROPANOL
3.   2210 0.37    118 C   5 H 10 O  3                            2-METHOXYETHYL ACETATE
4.    609 0.34     88 C   5 H 12 O  1                            ETHYL ISOPROPYL ETHER
5.    615 0.33     88 C   4 H  8 O  2                            BUTANOIC ACID (BUTYRIC ACID)

SPEC # 4221  ELUTION TIME=   25.78 MIN  ELUTION TEMP=   0  TIC MAX=   848

     ID#    R      MWT    FORMULA                                NAME
1. 14543 0.59     236 C   7 F  8                                 TRIFLUOROMETHYL-PENTAFLUOROBENZENE
2.  9749 0.35     186 C   6 F  6                                 HEXAFLUOROBENZENE
3. 14540 0.35     236 C   6 F  7 P  1                            PENTAFLUOROPHENYLDIFLUOROPHOSPHINE
4. 14395 0.17     234 C   6 CL 1 F  5 S  1                       PENTAFLUOROBENZENE SULPHENYL CHLORIDE
5. 13839 0.17     227 C   9 H 13 N  3 O  4                       2*DEOXYCYTIDINE

SPEC # 4224  ELUTION TIME=   26.13 MIN  ELUTION TEMP=   0  TIC MAX=  1568

     ID#    R      MWT    FORMULA                                NAME
1.    471 0.50     84 C   6 H 12                                 1-HEXENE
2.    477 0.46     84 C   6 H 12                                 2-METHYL-1-PENTENE
3.    475 0.46     84 C   6 H 12                                 METHYLCYCLOPENTANE
4.    462 0.44     84 C   6 H 12                                 CIS-2-HEXENE
5.    483 0.44     84 C   6 H 12                                 TRANS-2-HEXENE

SPEC # 4241  ELUTION TIME=   28,12 MIN  ELUTION TEMP=   0  TIC MAX=  1696

     ID#    R      MWT    FORMULA                                NAME
1.    354 0.69     78 C   6 H  6                                 2,4-HEXADIYNE
2.    353 0.68     78 C   6 H  6                                 1,5-HEXADIYNE
3.    352 0.68     78 C   6 H  6                                 1,5 HEXADIENE-3-YNE (DIVINYL ACETYLENE)
4.    351 0.68     78 C   6 H  6                                 BUTADIENYL ACETYLENE
5.    350 0.51     78 C   6 H  6                                 BENZENE

SPEC # 4247  ELUTION TIME=   28,82 MIN  ELUTION TEMP=   0  TIC MAX=   960

     ID#    R      MWT    FORMULA                                NAME
1.    471 0.54     84 C   6 H 12                                 1-HEXENE
2.    470 0.50     84 C   6 H 12                                 HEXENE-2
3.    445 0.48     84 C   5 H  8 O  1                            CYCLOPENTANONE
4.    453 0.48     84 C   5 H  8 O  1                            2-METHYLCYCLOBUTANONE
5.   1235 0.46    102 C   6 H 14 O  1                            1-HEXANOL

SPEC # 4256  ELUTION TIME=   29.87 MIN  ELUTION TEMP=   0  TIC MAX=  1136

     ID#    R      MWT    FORMULA                                NAME
1.    136 0.81     64 O   2 S  1                                 SULFUR DIOXIDE
2.    358 0.29     80 C   1 H  4 O  2 S  1                       METHANESULFINIC ACID
3.   2175 0.29    118 CL  2 O  1 S  1                            THIONYL CHLORIDE
4.   6579 0.24    159 C   2 F  3 N  1 O  2 S  1                  TRIFLUOROMETHANESULFINIC ACID ISOCYANATE
5.  11530 0.23    203 H   1 N  3 O  4 S  3                       2H-1,3,5,2,4,6-TRITHIATRIAZIN-1,1,3,3-TETROXIDE
```

TABLE 3. TYPICAL PRINTOUT FOR IDENTIFICATION OF COAL COMPONENTS (AFTER CLEANUP)

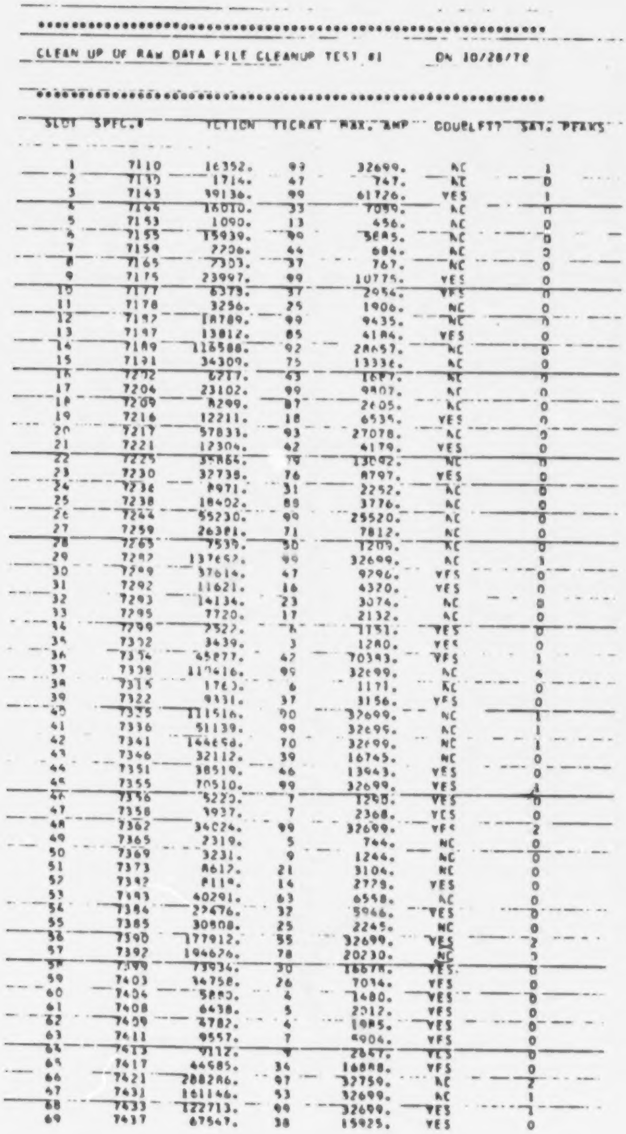
typical "before" and "after" pair of spectra from the coal run. Table 4 shows a comparison of a small selection of the data before and after CLEANUP. Complete accuracy is not claimed for this run, but CLEANUP improves the hit rate from nearly zero to an estimated 50%.

Several defects in the system emerged from this series of experiments. The most obvious defect involves situations in which strong peaks emerge from the column at sufficient amplitude to saturate the data acquisition system for a number of ions in several mass spectra. The most obvious case of this involves the component benzene (Figure 13, spectrum number ca. 13207), which eluted from the column as a very broad band extending for about 9 spectra (at 6 sec/scan this corresponds to a 54 second peak). This component showed saturated peak intensities for five major ions at m/z 50, 51, 52, 77, and 78. As a result of this situation, the ^{13}C isotope at m/e 79 was computed to have a relative intensity of 50%. As a result, the search program subsequently misidentified the major peaks in the spectrum and the search for benzene failed. This problem was compounded because, in addition, the large amount of benzene present caused other components, mainly from the "stationary" phase in the column to coelute, giving rise to a series of small peaks at high masses which the CLEANUP program dutifully processed and included in the unknown spectrum.

The problem of saturated peaks in spectrum identification is a serious one. Notwithstanding the fact that it is obviously wrong to try to identify components from a GC run when the amounts injected saturate the data acquisition system, these components are so strong and obvious that most investigators would be highly dissatisfied if simple compounds such as benzene were misidentified by a computer software system. There are several ways to remedy this situation. Currently, compounds with saturated peaks are labelled on the output printout (Figure 16) and thus warn the investigator of a potential problem. Also, a possible remedy for future development would be to raise cutoff limits on peak intensities when saturation was encountered so as to exclude weak peaks which otherwise would be accepted for searching.

Table 3. COMPARISON OF SELECTED SPECTRA FROM COAL GASIFICATION SAMPLE SEARCHED BEFORE AND AFTER CLEANUP

Before CLEANUP			After CLEANUP		
MWT	NAME	R	R	NAME	MWT
		Spectrum #13209 Elution Time = 24.38 TIC Max. 8081			
218	(AK-33)Methylcyclopentadienyl Maganese Tricarbonyl	0.19	0.21	Ethyl 2-Keto-2-Penylethandate	178
200	Alpha-(Bromoethyl)Benzyl Alcohol	0.16	0.20	Ethyl 2-Hydroxy-2-Phenylethandate	180
160	P-Phenylene Diisocyanate	0.15	0.20	P-Phenylene Diisocyanate	160
133	2-Methylphenyl Azide (D-Tolyl Azide)	0.15	0.20	Alpha-(Bromomethyl)Benzyl Alcohol	200
		Spectrum #13219 Elution Time = 25.55 TIC Max. 654			
198	7-Methyltridecane	0.15	0.43	2-pentanone	86
142	2,6-Dimethyl-3-Heptanone	0.14	0.42	Tri-Methylacetamide	101
121	3-Acetylpyridine	0.13	0.42	2-methyl-Pentanal	100
178	Ethyl 2-Keto-2-Phenylethandate	0.13	0.38	N-Butanal-D-Methyloxime	101
		Spectrum #13227 Elution Time = 26.48 TIC Max. 984			
130	Trichloroethylene	0.22	0.38	Trichloroethylene	130
132	1,1,1-Trifluoro-3-Chloropropane	0.16	0.23	1,1,1-Trifluoro-3-Chloropropane	132
146	N-Carbamoyl-2-Imino-1,3-Oxathiolane	0.15	0.22	1,2-Dichloro-1,2-Difluoroethylene	132
132	2-Thiedatane	0.15	0.22	1,1-Dichlorodifluoroethylene	132
		Spectrum #13233 Elution Time = 27.18 TIC Max. 862			
202	1,4-Dibutoxybutane	0.16	0.56	N-Heptane	100
132	1,3-Dimethoxy-2,2-Dimethylproprane	0.15	0.47	3-Methylhexane	100
162	1-(2-Butoxyethoxy)Ethanol	0.14	0.44	2,3-Dimethylpentane	100
158	2,2-Dimethyl-5-Methylene-3,7-Dioxa-Octane	0.14	0.42	2-Methyl-3-Pentanone	100
		Spectrum #13248 Elution Time = 28.93 TIC Max. 559			
98	Methylcyclohexane	0.28	0.75	Methylcyclohexane	98
98	Cycloheptane	0.24	0.74	Cycloheptane	98
112	1-Methyl-1-Ethylcyclopentane	0.22	0.60	Cyclohexanone	98
98	Cyclohexanone	0.22	0.58	2-Methylcyclopentandne	98
		Spectrum #13254 Elution Time = 29.63 TIC Max. 514			
146	2-Thianonane	0.21	0.40	2,2-Dimethylcyclobutanone	98
146	1-Octanethiol	0.19	0.37	2-Cyclohexene-1-ol	98
188	1-Undecanethiol	0.17	0.37	Cycloheptane	98
132	1-Heptanethiol (N-Heptyl) Mercaptan	0.16	0.36	3-Cyclohexenol	98



```
***********************************************************
CLEAN UP OF RAW DATA FILE CLEANUP TEST #1      ON 10/28/78
***********************************************************
 SLOT  SPEC.#      TCTION  TICRAT  MAX. AMP  DOUBLET?  SAT. PEAKS

    1    7110      16352.     97    32699.      NO          1
    2    7139       1714.     47      747.      NO          0
    3    7143      39136.     99    61726.     YES          1
    4    7144      16010.     33     7059.      NO          0
    5    7153       1090.     13      456.      NO          0
    6    7155      15939.     90     5885.      NO          0
    7    7159       2206.     44      684.      NO          0
    8    7165       7303.     37      767.      NO          0
    9    7175      23997.     99    10775.     YES          0
   10    7177       6379.     37     2954.     YES          0
   11    7178       3256.     25     1906.      NO          0
   12    7192      18789.     99     9435.      NO          0
   13    7197      13812.     85     4184.     YES          0
   14    7199     116588.     92    28657.      NO          0
   15    7191      34309.     75    13336.      NO          0
   16    7202       6717.     63     1687.      NO          0
   17    7204      23102.     99     9807.      NO          0
   18    7209       8299.     87     2605.      NO          0
   19    7216      12211.     18     6535.     YES          0
   20    7217      57833.     93    27078.      NO          0
   21    7221      12304.     42     4179.     YES          0
   22    7225      35864.     79    13092.      NO          0
   23    7230      32738.     76     8797.     YES          0
   24    7236       8971.     31     2252.      NO          0
   25    7238      18402.     89     3776.      NO          0
   26    7244      55230.     99    25520.      NO          0
   27    7259      26381.     71     7812.      NO          0
   28    7263       7539.     30     1209.      NO          0
   29    7287     137652.     99    32699.      NO          1
   30    7289      37614.     47     9796.     YES          0
   31    7292      11621.     16     4320.     YES          0
   32    7293      14134.     23     3074.      NO          0
   33    7295       7720.     17     2132.      NO          0
   34    7299       7522.      8     1151.     YES          0
   35    7302       3439.      3     1280.     YES          0
   36    7304      65877.     62    70393.     YES          1
   37    7309     119416.     99    32699.      NO          4
   38    7315       1763.      6     1171.      NO          0
   39    7322       9331.     37     3156.     YES          0
   40    7325     111516.     90    32699.      NO          1
   41    7336      51139.     99    32695.      NO          1
   42    7341     144658.     70    32699.      NO          1
   43    7346      32112.     39    16745.      NO          0
   44    7351      38519.     46    13943.     YES          0
   45    7355      70510.     99    32699.     YES          1
   46    7356       5220.      7     1250.     YES          0
   47    7358       3937.      7     2368.     YES          0
   48    7362      34024.     99    32699.     YES          2
   49    7365       2319.      5      744.      NO          0
   50    7369       3231.      9     1244.      NO          0
   51    7373       8612.     21     3104.      NO          0
   52    7382       8119.     14     2779.     YES          0
   53    7383      40291.     63     6558.      NO          0
   54    7384      22676.     32     5966.     YES          0
   55    7385      30808.     25     2245.      NO          0
   56    7390     177912.     55    32699.     YES          2
   57    7392     194676.     78    20230.      NO          0
   58    7399      73934.     30    16678.     YES          0
   59    7403      34758.     26     7034.     YES          0
   60    7404       5883.      4     1480.     YES          0
   61    7408       6438.      5     2012.     YES          0
   62    7409       3782.      4     1085.     YES          0
   63    7411       9557.      7     4904.     YES          0
   64    7413       9112.      9     2657.     YES          0
   65    7417      44585.     34    16888.     YES          0
   66    7421     288286.     97    32759.      NO          2
   67    7431     161146.     53    32699.      NO          1
   68    7433     122713.     99    32699.     YES          1
   69    7437      67547.     38    15925.     YES          0
```

Figure 16. Chemist's printout from CLEANUP.

```
 70    7439    27609.   19    5933.   NO     0
 71    7445    48464.   43   15673.   YES    0
 72    7447    14978.   11    4088.   YES    0
 73    7448    32380.   28    7289.   NO     0
 74    7453    12165.   10    1736.   NO     0
 75    7458     5804.    5    1529.   YES    0
 76    7459    15669.   15    3084.   YES    0
 77    7463    20917.   22    6761.   NO     0
 78    7468    80867.   99   32699.   YES    2
 79    7469     3856.    3    1986.   YES    0
 80    7471     5719.    6    1777.   YES    0
 81    7472    21524.   28    3765.   NO     0
 82    7478   259586.   74   32699.   NO     7
 83    7482    22038.   21    3463.   YES    0
 84    7483     7770.    9    1012.   YES    0
 85    7484     4966.    6     713.   YES    0
 86    7486     5504.    4    1393.   YES    0
 87    7489   124812.   67   18205.   NO     0
 88    7490   101103.   54   29322.   NO     0
 89    7497    19823.   22    2731.   YES    0
 90    7498    18524.   12    2777.   NO     0
 91    7504    13503.    6    1424.   YES    0
 92    7505    81709.   39   14600.   NO     0
 93    7511   307985.   97   32699.   NO     2
 94    7513    33757.   17    7284.   NO     0
 95    7516     2664.    3     514.   NO     0
 96    7520    48239.   42    9673.   NO     0
 97    7524    65144.   57    9421.   YES    0
 98    7528    18459.   27    4667.   NO     0
 99    7532     4643.    7     946.   NO     0
100    7534     5336.    9    1000.   NO     0
101    7544    17751.   29    2080.   NO     0
102    7546    14371.   25    3540.   NO     0
103    7552    26179.   49    5355.   NO     0
104    7556    94184.   53   20048.   YES    0
105    7557   100720.   99   22306.   NO     0
106    7559     2218.    3     540.   NO     0
107    7562    79764.   88   15105.   NO     0
108    7564    13380.   23    3684.   NO     0
109    7569    57043.   79   13283.   YES    0
110    7571    20409.   31    4879.   YES    0
111    7573    29182.   42    6078.   YES    0
112    7577    10942.   17    3371.   NO     0
113    7588    27725.   40    6134.   YES    0
114    7590    38307.   56    9176.   NO     0
115    7596    97445.   91   22153.   NO     0
116    7602     3965.    6    1505.   NO     0
117    7606     2180.    3    1642.   NO     0
118    7615     6393.   10    1785.   NO     0
119    7619     3496.    5    1390.   YES    0
120    7623    16928.   24    2310.   NO     0
121    7629     3851.    6     771.   YES    0
122    7630     6274.   10     700.   YES    0
123    7634    74127.   56   16870.   NO     0
124    7637     3056.    4     772.   YES    0
125    7639     4655.    7    1607.   NO     0
126    7649     3291.    6     890.   NO     0
127    7659     4311.    5    1597.   YES    0
128    7660    21496.   34    8647.   NO     0
129    7672    38809.   99   32699.   NO     2
130    7674    40638.   48   15489.   YES    0
131    7675     1339.    2    1022.   NO     0
132    7696    19246.   34    6373.   NO     0
133    7703     3835.    8    1041.   NO     0
134    7705     1525.    3     615.   NO     0
135    7710     1563.    3    1228.   NO     0
136    7717    36074.   64    8402.   NO     0
137    7723    29546.   53    7979.   NO     0
138    7725    10576.   26    1758.   NO     0
```

Figure 16, Continued.

Analysis of Environmental Samples

As a further test of the proper operation of the CLEANUP program, a series of runs were processed at RTI using data which had been previously analyzed by hand. Seven runs were processed which represented a selection of different types of samples which possessed a variety of functional group types. These runs are briefly described in Table 5. Of all of the runs for which data were available, one represented a very complex mixture containing a wide variety of components and was chosen for extended study. The sample was that derived from ambient air in a basement of a residence which had been built over an industrial chemical dump and which had been perfused with a large variety of organic components. The air sample when analyzed by GC/MS, showed the presence of in excess of 100 components. A total ion current plot of this run is shown in Figure 17. The sample was first processed by the CLEANUP program using the best available values for the input parameters, and the results compared with the manual identification. The results of this comparison are shown in Table 6.

Notwithstanding a small number of errors, it is obvious that the CLEANUP program is operating very successfully, and is capable of extracting correct spectra from mixtures and identifying them even in very difficult and complex mixtures. The table shows a number of interesting results. It should be noted that in the majority of cases, the manual identification and the computer results were the same. When this is combined with a high correlation coefficient ($R>0.6$), then the identification can be said to be correct and confirmed. In such cases the table entry is marked with an H (hit).

In some cases the computer results identified a component which is known to have a spectrum very similar to the manually identified compound. The correct identity often can be deduced by inspection, e.g. toluene might be identified as a methyl hexatriyne, or vinylidine chloride confused with dichloroethylene. This category comprises mainly isomers or homologues. These cases were also marked with an H. When the computer printed the manually identified compound as a 2nd to 5th choice, this number was also included in the table under the column heading, "Found".

TABLE 5. DESCRIPTION OF TYPES OF SAMPLES INVESTIGATED TO TEST OUT CLEANUP

All samples were run by desorption from TENAX GC cartridges.

Type of Sample	Contents
1. UNC smog chamber	esters nitrates dioxanes nitriles aldehydes hydrocarbons low m.w. alkyl aromatics ketones
2. Oil shale volatiles	nitriles alcohols sulfur compounds ketones pyridines pyrazoles quinolines hydrocarbons aromatics
3. Volatiles from in situ coal gasification (tar fraction)	alkanes alkyl aromatics very few hetero atom containing compounds
4. Volatiles from in situ coal gasification (water fraction)	alkanes sulfur compounds pyridines oxygenated hydrocarbons
5. Organics from air in basements	alkanes aromatics chlorines, bromine and fluorine containing compounds
6. Organics in air in the vicinity of a Du Pont plant	halogenated hydrocarbons aldehydes ketones phenols nitro aromatics
7. Organics near a pesticide plant	halogenated hydrocarbons sulfur compounds oxygenated hydrocarbons (aldehydes, ketones, phenols, esters)

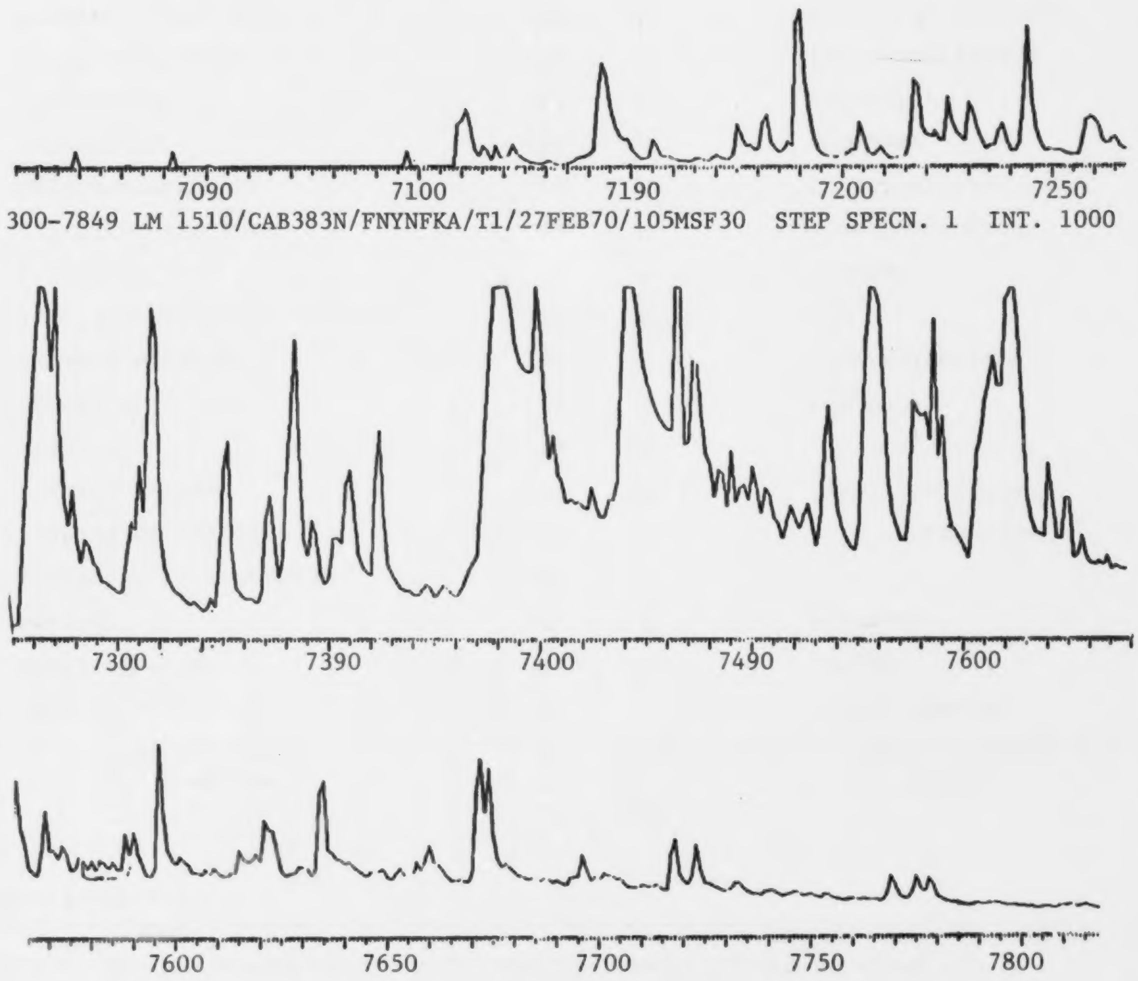


Figure 17. Total ion current plot from test run.

TABLE 6. COMPARISON OF MANUAL VS. COMPUTER IDENTIFICATION OF GC-MS RUN

Run #: 1510-CAB-3B3W-FNYNFKB

SP #	Retention Time	Compound Identification	R	Hit or Miss	Found	Manual Identification
7110	12.72	Dimethylamine	.84	M	5	Carbon dioxide
7130	15.05	Vinyl fluoride	.63	H		C_3H_8O?
7143	16.57	Acetone	.89	H		Acetone
7144	16.68	Fluorotrichloromethane	.85	H		--
7153	17.73	Vinylidine chloride	.78	H		Dichloroethylene
7155	17.97	Dichloromethane	.96	H		Dichloromethane
7159	18.43	Freon 113	.54	H		--
7165	19.13	Acetic acid	.71	M		C_4H_8O?
7175	20.30	2-Methylpentane	.80	H		2-Methylpentane
7177	20.55	1,2-Dimethylpropyl acetate	.55	M		C_5H_{10}
7178	20.65	n-Butanal	.75	H		Butanal
7182	21.12	2,2-Dimethylbutane	.69	H		3-methylpentane
7187	21.70	Hexafluorobenzene	.78	H		Hexafluorobenzene
7189	21.93	n-Hexane	.82	H		n-Hexane
7191	22.17	Chloroform	.63	H		Chloroform
7202	23.45	Perfluorotoluene	.71	H		Perfluorotoluene
7204	23.68	Methylcyclopentane	.77	H		Methylcyclopentane
7209	24.27	1,1,1-Trichloroethane	.68	H		1,1,1-Trichloroethane
7216	25.08	Isopropyl acetate	.76	H		Isopropyl acetate

(continued)

TABLE 6 (continued)

SP #	Retention Time	Compound Identification	R	Hit or Miss	Found	Manual Identification
7218	25.20	Benzene	.90	H		Benzene
7220	25.67	Acetic acid	.67	H		Acetic acid
7225	26.13	2-Methylhexane	.78	H		2-Methylhexane
7230	26.72	3-Methylhexane	.89	H		3-Methylhexane
7235	27.42	Dimethylcyclopentane	.82	H		C_9H_{14}
7238	27.65	Trichloroethylene	.75	H		Trichloroethylene
7244	28.35	*n*-Heptane	.97	H		*n*-Heptane
7259	30.10	Methyl isobutyl ketone	.54	H		4-Methyl-2-pentanone
7265	30.80	*n*-Heptyl formate	.76	?		C_8H_{18}
7281	32.78	1,5-Heptadien-3-yne	.60	H	5	Toluene
7289	33.60	3-methylheptane	.91	H		C_8H_{18}
7292	33.95	Hexanal	.61	H		$C_6H_{12}O$?
7293	34.07	1,3-Dimethylcyclohexane	.62	H		C_8H_{16}
7295	34.30	Ethyl *n*-butanoate	.52	H		--
7300	34.77	4-Methyltriazole	.46	?		C_8H_{16}
7303	35.12	*n*-Butyl acetate	.42	H		*n*-Butyl acetate
7305	35.35	3,4-Epoxy-2-hexanone	.52	M		*n*-Octane
7309	35.82	Tetrachloroethylene	.58	H		Tetrachloroethylene
7315	36.63	Hexamethylenecyclohexasiloxane	.57	?		--
7322	37.45	2,6-Dimethylheptane	.78	H		C_9H_{20}
7325	37.80	Chlorobenzene	.80	H		Chlorobenzene

(continued)

TABLE 6 (continued)

SP #	Retention Time	Compound Identification	R	Hit or Miss	Found	Manual Identification
7336	39.08	1,3-Dimethylbenzene	.83	H		Ethylbenzene
7342	39.67	N-Methyl-n-pentylhydrazine	.74	M	3	Xylene isomer
7346	40.25	3-Methyloctane	.81	H		C_9H_{20}
7351	40.83	Benzocyclobutane	.75	H	3	Styrene
7355	41.30	p-Xylene	.77	H	3	o-Xylene
7356	41.42	1,3-Dimethylcyclopentane	.40	?		--
7358	41.65	1-Methyl-4-ethylcyclohexane	.43	H		C_9H_{18}
7361	42.16	n-Nonane	.69	H		n-Nonane
7365	42.47	Diethyl methylvinylsilane	.51	?		--
7369	42.93	1-Methylethylcyclohexane	.75	H		C_9H_{18}
7373	43.40	Isopropylbenzene	.56	H		Isopropylbenzene
7381	44.45	Cyclohexyl acrylate	.63	?		C_3-Alkylcyclohexane
7383	44.57	3-Methylnonane	.52	H		$C_{10}H_{22}$
7384	44.68	Di(2-ethylhexyl) ether	.56	H		--
7385	44.80	2-Methyl-5-ethyl heptane	.28	H		$C_{10}H_{22}$
7390	43.38	Benzyl chloride	.58	H	2	Chlorotoluene
7392	45.62	4-Azido-3-nitrotoluene	.37	M		--
7399	46.43	Isopropylbenzene	.38	H		C_3-Alkylbenzene
7403	46.90	3-Methylnonane	.77	H		$C_{10}H_{22}$
7404	47.02	3-Phenyl-3-methylazetidine	.66	?		--
7408	47.48	p-Menthan-9-ol	.45	H	2	$C_{10}H_{20}$

(continued)

TABLE 6 (continued)

SP #	Retention Time	Compound Identification	R	Hit or Miss	Found	Manual Identification
7409	47.60	β-Thujene	.41	H		$C_{10}H_{20}$
7412	47.83	Valeric acid	.66	H	3	Heptanoic acid?
7413	48.07	4-*n*-Propyl-3-heptene	.42	M		C_3-Alkylbenzene
7417	48.53	Octamethylcyclooctasiloxane	.52	?		--
7420	49.00	o-Chlorobenzoyl chloride	.43	H	2	Dichlorobenzene isomer
7431	50.17	1,3-Dichlorobenzene	.53	?		C_4-Alkylbenzene
7433	50.40	Cuminic aldehyde	.48	?		C_3-Alkylbenzene
7438	50.87	1-*p*-Menthadiene	.76	H		$C_{11} + C_{10}$ isomer
7439	51.10	Isobutylcyclohexane	.51	M		Bromotoluene isomer
7444	51.80	*p*-Methylacetophenone	.68	H	4	C_4-Alkylbenzene
7447	52.03	4-Ethylheptane	.53	H	2	$C_{11}H_{24}$ Isomer
7448	52.15	O-Decyl-hydroxylamine	.50	?		$C_{11}H_{24}$ Isomer
7453	52.73	Bicyclo(4.4.0)decane	.42	?		$C_{11}H_{24}$
7458	53.32	3,5-dimethylmethyl benzoate	.33	M		C_4-Alkylbenzene
7459	53.43	Ar,α-dimethylstyrene	.50	H		--
7463	53.90	Linalol	.39	?		$C_{11}H_{22}$
7468	54.48	*n*-Undecane	.87	H		*n*-Undecane
7469	54.60	*t*-Butylbenzene	.28	?		--
7471	54.83	Methyl benzoylacetylene	.36	M		Tetrachlorobenzene
7472	54.95	o,α-Dichlorotoluene	.38	H		Trichlorotoluene
7478	55.65	2,6-Dichlorotoluene	.70	H		Dichlorotoluene

(continued)

TABLE 6 (continued)

SP #	Retention Time	Compound Identification	R	Hit or Miss	Found	Manual Identification
7482	56.12	α-Neoisomenthol	.21	M		Pentachlorobutadiene
7484	56.23	Bicyclo(3.3.1)nonane-2-ol	.28	M		Trichlorobenzene
7485	56.35	Methyldichlorocyclopentadiene	.25	M	2	Trichlorobenzene
7486	56.58	3,4-Dimethylstyrene	.45	?		C_5-Alkylbenzene
7488	56.82	α-Pinene oxide	.77	H		$C_4H_{16}O$
7490	57.05	2,6-Dichlotoluene	.85	H		Dichlorotoluene
7497	57.87	N-Acetyl-phenylalanine	.21	M		--
7498	57.98	2-Methylnonane	.43	H		--
7504	58.68	1,2,4-Trichlorobenzene	.43	H		Trichlorobenzene
7506	58.80	1,2,4-trichlorobenzene	.48	H		Trichlorobenzene
7511	59.50	Quinoline butiodide	.43	M	3	Naphthalene
7513	59.73	n-Dodecane	.65	?		Methyl salicylate
7515	60.08	Verbenone	.33	?		$C_{12}H_{26}$
7520	60.55	1,3,5-Trichlorobenzene	.72	H		Trichlorobenzene
7525	61.02	Hexachlorobutadiene	.69	H		Hexachlorobutadiene
7528	61.48	Carvone	.74	H		?
7532	61.95	Pentobarbital-TMS ether	.20	M		Unsat. hydrocarbon
7534	62.18	Cyclohexylmethanol	.31	M		$C_{13}H_{26}$
7544	63.35	α,2,4-trichlorotoluene	.37	H		Trichlorotoluene
7546	63.58	2-Methyl-5-ethylheptane	.64	?		$C_{14}H_{28}$

(continued)

TABLE 6 (continued)

SP #	Retention Time	Compound Identification	R	Hit or Miss	Found	Manual Identification
7552	64.28	Bornyl formate	.59	H		$C_{10}H_{16}$(?)
7556	64.75	n- Tridecane	.67	H		n-Tridecane
7558	64.87	α,3,4-Trichlorotoluene	.67	H		Trichlorotoluene
7559	65.10	Haloperidol	.28	M		Methylnaphthalene
7561	65.45	α,3,4-Trichlorotoluene	.25	H		Trichlorotoluene
7564	65.68	1-Methylnaphthalene	.61	H		α-Methylnaphthalene
7568	66.27	Chlorophenyldibenzoazepine	.32	M		$C_{14}H_{30}$?
7571	66.50	1,2,3,4-Tetrachlorobenzene	.68	H		Tetrachlorobenzene
7573	66.73	α,2,4-Trichlorotoluene	.63	H		Trichlorotoluene
7577	67.20	2,5-Dimethylundecane	.48	H		$C_{13}H_{26}$ isomer
7588	68.48	Acenaphthene	.50	H	2	Biphenyl
7590	68.72	1,2,3,4-Tetrachlorobenzene	.68	H		Tetrachlorobenzene
7596	69.42	n-Tetradecane	.79	H		n-Tetradecane
7602	70.12	2,6-Dimethylnaphthalene	.47	H		C_2-Alkylnaphthalene
7607	70.58	2-Amino-4-hydroxupteridine	.40	M		C_2-alkylnaphthalene
7615	71.63	α-murolene	.24	M		Alkylbenzene?
7619	72.10	1-Cyclohexyloctane	.39	H	2	$C_{16}H_{32}$
7623	72.57	2,6-Di-t-butyl-1,4-benzoquinone	.61	H		$C_6H_{24}O$?
7624	73.27	7-Acetoxy-p-menth-1-en-3-one	.31	?		--
7630	73.38	2,6-Dichlorobenzal chloride	.32	H	2	Tetrachlorotoluene
7634	73.85	n-Pentadecane	.82	H		n-Pentadecane
7637	74.20	Methyl laurate	.38	?		$C_{15}H_{30}$

(continued)

TABLE 6 (continued)

SP #	Retention Time	Compound Identification	R	Hit or Miss	Found	Manual Identification
7639	74.43	Trihydroxy benzoic acid, TMS	.19	M		C_3-Alkylnaphthalene
7649	75.60	Pentachlorobenzene	.38	?		C_5-Alkylbenzene
7659	76.77	Dicyclohexyl ether	.30	?		--
7660	76.88	Diethyl phthalate	.68	H		Diethyl phthalate
7672	78.28	Neopentylphosphonyldichloride	.41	M		Trimethylpentadiol diacetate
7674	78.52	n-Hexadecane	.58	H		n-Hexadecane
7675	78.63	2-Phenyldecane	.39	H		Sat. hydrocarbon
7696	81.08	Pristane	.66	H		$C_{18}H_{38}$
7703	81.90	Dimethylbenzylisobutyl ether	.23	M		$C_{16}H_{32}$
7705	82.13	Methylfuranyl THP-ether	.29	M		--
7710	82.72	Trimethylamine	.65	M		--
7718	83.53	n-Tetradecane	.77	H	3	n-Heptadecane
7723	84.23	Pristane	.82	H		$C_{19}H_{40}$
7775	90.30	Myristic acid	.48	H		?

An interesting situation occurred in a few cases in which manual identification was not possible or uncertain. In those cases the manual identification column contains a question mark. If in addition the R factor from the computer is sufficiently high, it is reasonable to assume that the computer identification is correct and the manual identification is incorrect. This is particularly true with closely spaced doublets where frequently it is nearly impossible to separate components manually. In these cases, the computer identification is more certain than the human one. When this occurred, a question mark or even an H was entered in the hit column. Components missed in the manual interpretation were marked with a dash in the appropriate column.

In a small number of cases, computer identification is undoubtedly wrong. This is usually accompanied by low R factors ($R<0.3$). Also when peaks are weak or small components of multiplets, it is difficult to identify them by computer. Such cases are marked with an M in the hit column ("Miss").

Comparison of figures clearly shows that the CLEANUP program materially simplifies the spectra prior to their submission to the search program. The data here presented show that by combining CLEANUP with an efficient search procedure, results can be obtained which are comparable with human interpretation. In summary, of 138 components, definite hits were obtained on 90 components, clear misses occurred in 26 cases, with 22 cases unclear. Removing the questionable category, the system recorded 78% hits and 22% misses, a rather remarkable performance for this very complex and convoluted mixture.

CLEANUP CONTROL PARAMETERS

The conclusion of Phase II consisted of variation of the CLEANUP control parameters in order to determine their effect on the output spectra and to arrive at an optimum setting for each. The fifth of the eight runs described above was chosen as a test case. Most of the parameters represent threshold settings concerned with eluant detection: peak sharpness requirements, total ion current threshold, closest expected distance between neighboring eluants, the percent of each peak attributable to background, etc. Two parameters are simply switches controlling optional diagnostic messages useful for program debugging.

An attempt was made to determine optimal parameter settings by varying parameters in a regular manner and analyzing the results using the SPSS statistical package. This approach was not fruitful since there was no obvious demarcation between what constitutes reak peaks and noise peaks, and so the statistics did not show a clear preference for one setting of the parameters, but only a contininuum of change. Trial and error variation of the parameters, though less sophisticated, proved more effective. The general procedure was to vary each parameter in both directions about the value suggested by the program authors in the original CLEANUP documentation. A summary of conclusions and recommended values is shown in Table 7. The two things which most aided determining the effects of parameter variation were the "chemist's summary printout" of CLEANUP (Figure 17) and the output of the RTI search program (Table 4). It is to be emphasized that these values represent optimum settings for a VARIAN CH-7. Our suggested values for the INCOS test data are shown in Figure 27.

The chemist's printout is a summary list of components found by CLEANUP, along with information regarding the presence of doublets and saturated peaks. By comparing the change of the number of eluants indicated in the chemist's printout together with the most likely candidates for those eluants found by the RTI search program, it was possible to determine when erroneous peaks were beginning to be picked up by CLEANUP.

Following is a technical review of the analysis of parameters, primarily of interest to programming specialists.

The first two variables control the print options of the CLEANUP program and do not affect any internal calculations. Either variables can be set to 0 (do not print) or 1 (print). The first option (IPFLAG) determines whether or not the chemist's printout is to be printed. The listing of all eluants found by CLEANUP takes relatively little space, and has proved helpful in almost all test runs. We feel that the chemist's printout should always be requested. The second option (IDEBUG) determines whether or not intermediate calculations and decisions inside CLEANUP are to be printed. Information printed includes histogram

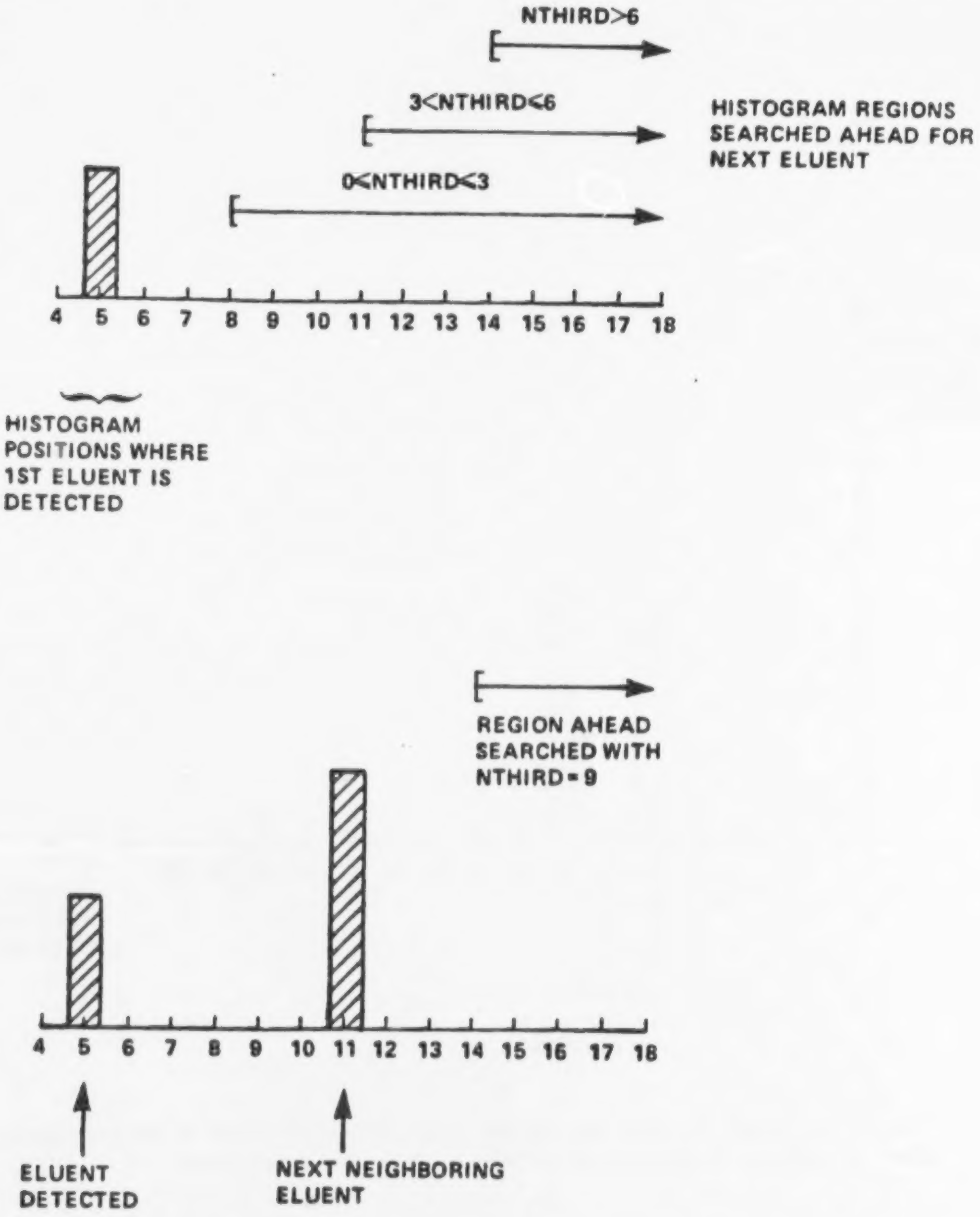
TABLE 7. CLEANUP INPUT PARAMETER SUMMARY

Name	Function	Possible values	Suggested values
IPFLAG	Determine whether or not to generate chemist's printout	1 = Generate printout 0 = Suppress printout	1 = Chemist's printout is helpful in all cases
IDEBUG	Determine whether or not to generate debugging output	1 = Generate printout 0 = Suppress printout	0 = Printout necessary only when looking for errors
NTHIRDS	Set minimum # of scans for expected doublets	3,6,9 in units of 1/3 scan	3 = Expected distance as short possible
ITOM	Set neighborhood in histogram window for eluent detection	1,2 units of 1/3 scans	2 = Widest setting
MINNY	Set minimum histogram ion current for 1 to 4 peak eluents	Any integer greater than or euqal to 0	1000 currently used, setting greater than 0 recommended
MINTIC	Set minimum histogram ion current for eluents with more than 5 peaks	Any integer greater than or equal to 0	1000 currently used and recommended, critical that setting greater than zero is used
RATM	Minimum rate for eluent detection	Any real number greater than 0.0	0.0 = No minimum cutoff for peak rates appears necessary at this time
RATMXX	Eliminate column bleed from detected eluents by linear error ratio test	0.0 to 1.0	= .1, current value is .0756. It is critical that this parameter not be 0.0
ILEFT	Set number of scans required for increasing ion current prior to prospective model peak mode	1,2	2 = As wide a neighborhood about the mode as possible is recommended to ensure model accuracy
IRIGHT	Set number of scans required for decreasing ion current after prospective model mode	1,2,3	3 = Same as above
SATVAL	Tells CLEANUP instrument saturation value	All integers	= 200 below instrument saturation value (see text)

printouts, complete peak models, and decisions made as to whether the peak encountered was dismissed as noise, qualified as a singlet eluant, or whether the doublet resolver was invoked. With an average of 1 page of printout per eluant detected, the full printout for some test runs has frequently exceeded 200 pages. Although this information proved very helpful in finding some bugs in CLEANUP, we feel that unless thre is a specific need, this diagnostic output should be suppressed.

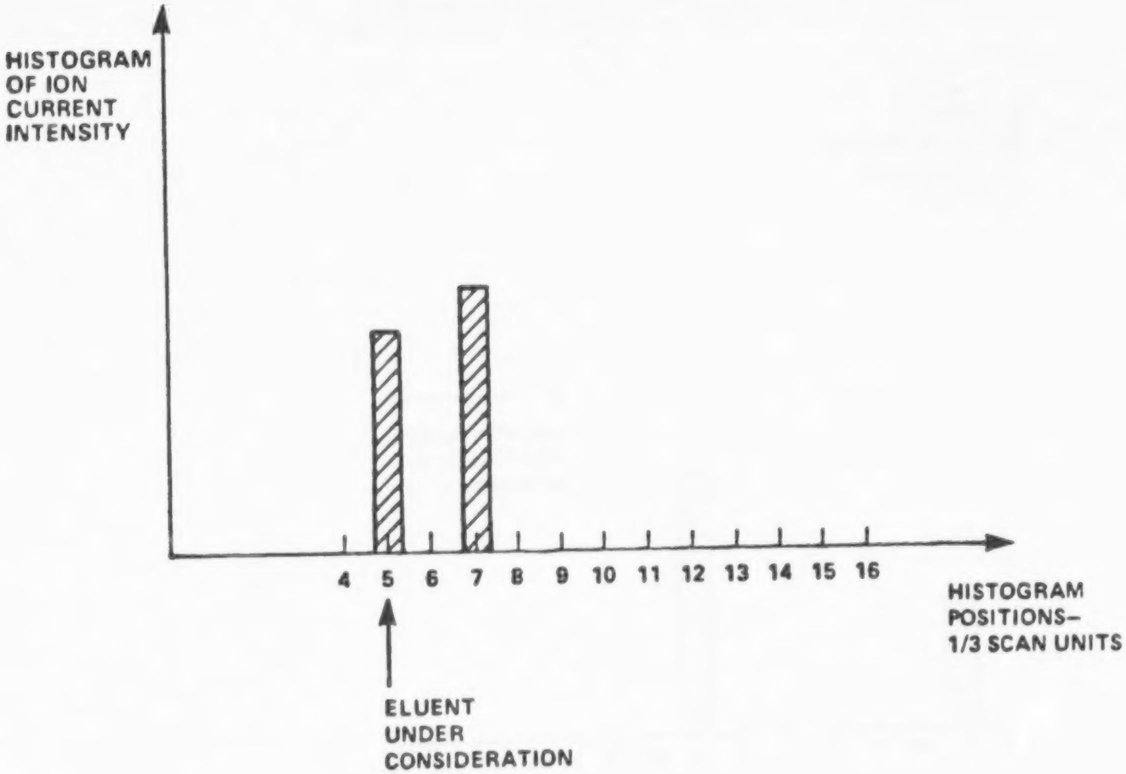
The first of the parameters affecting output spectra (NTHIRD) determines the closest distance which CLEANUP expects between neighboring eluants. Once the CLEANUP program detects an eluant from histograms of the ion current, it looks ahead in the histogram to determine if another eluant is in the vicinity just ahead. Although the variable NTHIRD is given in units of thirds of scans, only three values make a difference in terms of CLEANUP internal calculations. A value for NTHIRD of 3 means that neighboring eluants are expected no closer than 1 scan away, 6 means no closer than 2 scans, and 9 means that neighboring eluants are expected no closer than 3 scans from each other. The function of NTHIRD is shown graphically in Figure 18. From a theoretical point of view, there is no reason why this parameter should not be set to 3, setting the CLEANUP program to look for doublets within the finest possible grid. The danger of using a coarser setting is that some close pairs of eluants could be interpreted by the CLEANUP program as singlets, resulting in incorrect peak model choices and the deletion of unknowns. Our experience with varying this parameter over the test run however, showed only a mild dependence of results on the expected minimum distance, with the number of eluants decreasing by only 5% from the finest to the coarsest possible setting.

The detection window size (ITOM), has two possible values; 1 and 2, corresponding to the number of histogram positions which must be less than the central maximum in order for an eluant to qualify for detection. This is shown graphically in Figure 19. Currently, the most stringent setting is being used, requiring that histograms of ion currents for 2 members on either side of the central position be lower than the central value in order for an eluant to be detected. Statistics on this


NTHIRD>6
3<NTHIRD≤6
0<NTHIRD≤3
HISTOGRAM REGIONS SEARCHED AHEAD FOR NEXT ELUENT
4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
HISTOGRAM POSITIONS WHERE 1ST ELUENT IS DETECTED
REGION AHEAD SEARCHED WITH NTHIRD=9
4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
ELUENT DETECTED
NEXT NEIGHBORING ELUENT

Example of Eluent Missed Due to Incorrect Setting of NTHIRD Parameter

Figure 18. Graphic representation of possible settings of NTHIRD.



The ion at position 5 would be rejected with ITOM = 2 because of the neighboring eluent at position 7, whereas with ITOM = 1, it would be accepted.

Figure 19. Function of ITOM histogram window setting.

parameter were obtained, but in view of the fact that the CLEANUP algorithm does not try to resolve doublets closer than 1 scan apart and breaks down entirely for a distance of 1/3 scan, it was decided to keep ITOM at its most rigorous value, 2.

The input parameters MINNY and MINTIC are minimum TIC histogram intensity thresholds for eluant detection. It is important to note that these ion currents are tested after background noise is subtracted. There are two threshold settings, one for spectra containing very few ions, (e.g. coronene), and one for spectra containing very few ions, (e.g. coronene), and one for spectra containing 5 or more ions. In addition, the reconstructed spectrum is required to have a total ion current equal or greater than MINTIC otherwise it is not written to the output file or noted in the chemist's printout. TIC threshold requirement was tested first by setting both parameters to 0, effectively eliminating the TIC minimum cutoff. The result of this test was to more than double the number of eluants detected by CLEANUP. Many of these eluants when processed by the RTI search program were found to have less than 5 compounds in the entire library which passed its two out of six strongest peaks presearch test, indicating that the peaks consisted mainly of noise. It is thus concluded that some TIC requirement is necessary. We now believe that the value of 1000, presently used, is too high. The optimum setting of this threshold must be determined individually for every data system.

The parameter (RATM) sets a minimum sharpness on each model peak before it can qualify as an eluant. The suggested value is 0, meaning that no eluents are rejected, sharpness being positive definite. This parameter was left at 0 for two reasons. First, there was no evidence of spurious eluants being introduced due to insufficient sharpness, and second, the distribution of sharpness values for the test run gave no clues as to a best setting.

The function of the linear error ratio (RATMXX) is to remove those single ion traces (fragmentograms) which contain column bleed. In order to do this, a least squares fit to a straight line is calculated for an 11 scan window (5 scans on each side scan under consideration) and then the ratio of the root mean square deviation of the peak shape about the line to the average value of the line is calculated (Figure 20). As indicated, peak shapes which are "too flat" can be eliminated by the error ratio criterion. Another useful property of the error ratio is that small peaks against very limited background noise will still be considered as eluants (b), whereas the same small ripple against a strong background would be dismissed as noise (c). We have varied the minimum error ratio requirement and we have demonstrated that the correct assignment of this variable is critical. The minimum error ratio requirement was reduced from the suggested .0756 to .002. Under these conditions CLEANUP found the same number of peaks as it had with the higher value found (the error ratio test is applied after an eluant has been detected in order to determine if reconstruction is worthwhile), but the CPU time required by the search program doubled and recognition performance markedly deteriorated. Extraneous peaks were admitted by CLEANUP, causing a large fraction of the unknowns to pass the presearch cycle of SEARCH and also causing misidentification in the search program due to the presence of those peaks.

The parameter ILEFT sets the number of scans which must show an increasing ion current prior to the mode in order for an ion fragmentogram to be considered as a model peak. Correspondingly, IRIGHT sets the number of scans required to show a decreasing ion current after the mode for model consideration. We feel that ILEFT and IRIGHT should be left at their maximum permissable values, 2 scans prior to and 3 scans after the mode. The CLEANUP algorithm achieves maximum efficiency when a typical singlet peak is 5 to 10 scans wide. If the scan rate is correctly "tuned" the most stringent setting for ILEFT and IRIGHT ensures that well-behaved individual ion traces are used as model peaks.

The parameter SATVAL defines the saturation value used by the program in order to determine when peaks are saturated and need to be reconstructed via a ratio approach. Although this number would be expected to be the theoretical maximum for the A/D convertor used in the data acquisition system, we have found it necessary to set SATVAL to a value slightly below the theoretical maximum. Thus, for the VARIAN

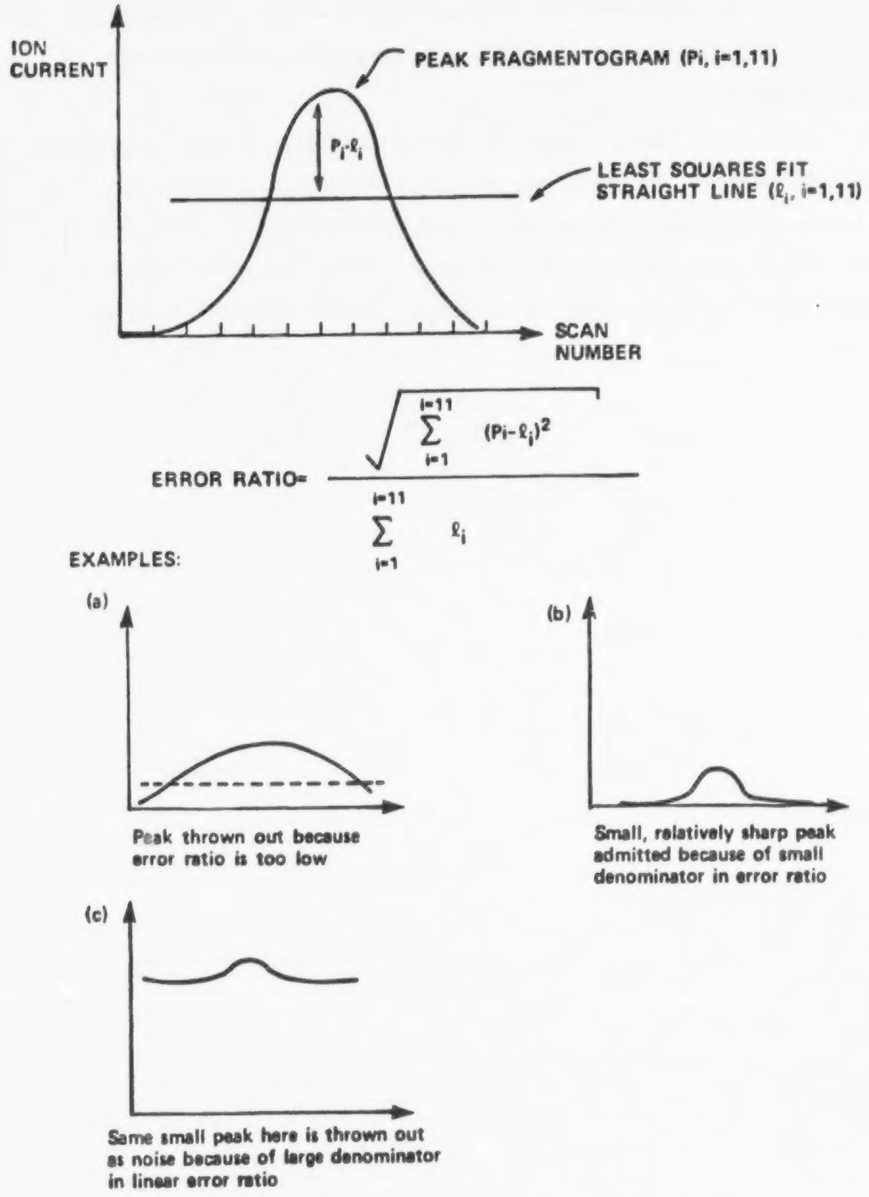
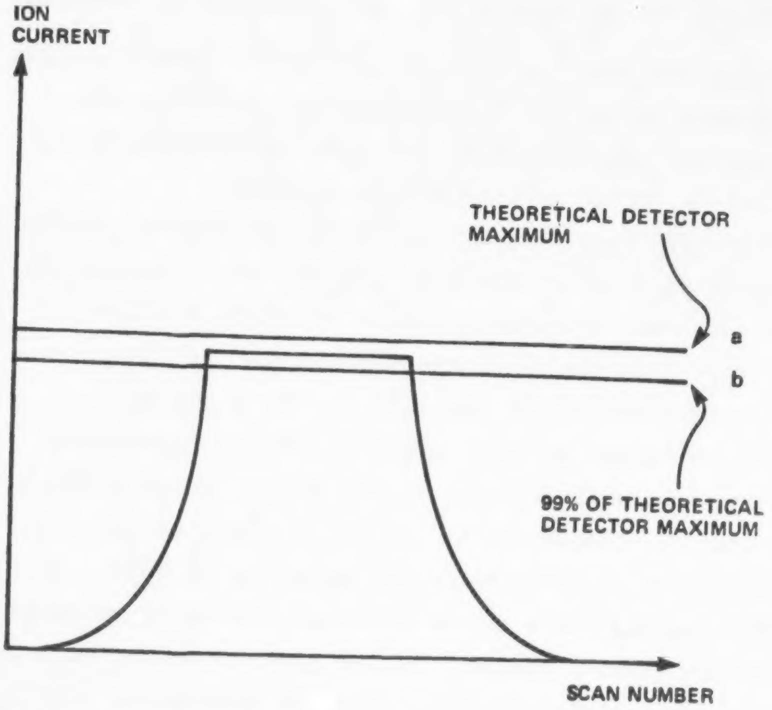

ION CURRENT
PEAK FRAGMENTOGRAM (Pi, i=1,11)
$P_i-\ell_i$
LEAST SQUARES FIT STRAIGHT LINE (ℓ_i, i=1,11)
SCAN NUMBER
ERROR RATIO= $\dfrac{\sqrt{\sum_{i=1}^{i=11} (P_i-\ell_i)^2}}{\sum_{i=1}^{i=11} \ell_i}$
EXAMPLES:
(a)
Peak thrown out because error ratio is too low
(b)
Small, relatively sharp peak admitted because of small denominator in error ratio
(c)
Same small peak here is thrown out as noise because of large denominator in linear error ratio

Figure 20. Effects of linear error ratio.

system, the maximum A/D output is 32,767 (2^{15}-1), we currently have SATVAL set to 32,000. For unknown reasons, A/D convertors often register values less than their maximum for known, saturated peaks. In such cases, saturated peaks can inadvertently be eliminated entirely as background noise (Figure 21).

Again, the values for the parameters shown in Table 7 are satisfactory for VARIAN CH-7 data. It must be realized however, that these values are instrument and data dependent, and optimum values for any particular data system must be determined empirically by adjustment of parameters, which will require some experience gained only with time.


ION CURRENT
THEORETICAL DETECTOR MAXIMUM
a
b
99% OF THEORETICAL DETECTOR MAXIMUM
SCAN NUMBER

With SATVAL setting at value (a), the saturated peak is constructed by CLEANUP to be entirely background noise, but with a slightly reduced setting (b) for SATVAL, the peak is recognized as saturated and handled by the correct routines.

Figure 21. Effect of SATVAL on CLEANUP.

SECTION 5

IMPLEMENTATION OF THE CLEANUP PROGRAM ON THE UNIVAC COMPUTER

Phase III of this contract consisted of making the CLEANUP program developed at RTI operational on the UNIVAC computer located at EPA. This work was carried out simultaneously with some improvements being made in the organization of the data formats in CLEANUP.

Because RTI's and EPA's environment for using the CLEANUP program are similar, it was decided to use the same CLEANUP source version for each. The general plan was to change the READ and WRITE statements inside CLEANUP to subroutine calls and then write a set of I/O subroutines for EPA and later a set of I/O routines for RTI's system.

There were two advantages to this approach. First, the CLEANUP program would require substantially less modification. In addition, if any bugs were to occur in CLEANUP in the future, it would be possible to investigate the problem by carrying out a debugging run at TUCC. This would verify whether the bug was a system oriented problem or an algorithm deficiency inside CLEANUP proper. If the problem was indeed a bug in CLEANUP, it could be corrected at TUCC. Systems problems at the UNIVAC would be corrected there. Modules required for these programs are shown on Figure 22. The code to be delivered to EPA consists of sections 1, 2 and 3. A short description of each of these modules follows.

Some minor modifications to the original CLEANUP program besides changing READ and WRITE statements were made. The most important of these was to expand the CLEANUP program so that it could process multiple runs on the same tape data set. To include this facility, the routines PARDIR (reading in the CLEANUP input parameters) and SETRUG (initializing the 14 spectra window with the first 13 spectra) had to be repositioned and rewritten.

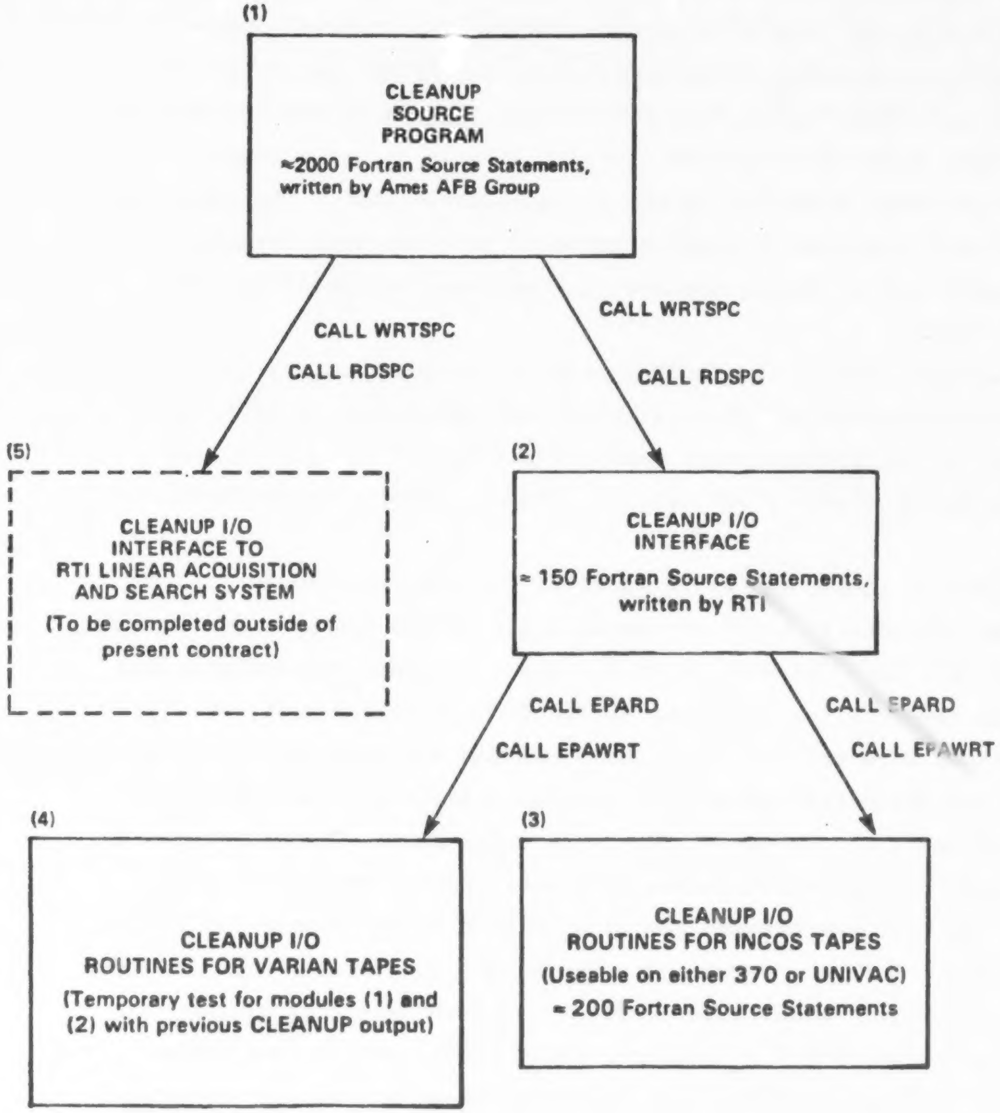

(1)
CLEANUP
SOURCE
PROGRAM
≈2000 Fortran Source Statements,
written by Ames AFB Group
CALL WRTSPC
CALL RDSPC
CALL WRTSPC
CALL RDSPC
(5)
CLEANUP I/O
INTERFACE TO
RTI LINEAR ACQUISITION
AND SEARCH SYSTEM
(To be completed outside of
present contract)
(2)
CLEANUP I/O
INTERFACE
≈ 150 Fortran Source Statements,
written by RTI
CALL EPARD
CALL EPAWRT
CALL EPARD
CALL EPAWRT
(4)
CLEANUP I/O
ROUTINES FOR VARIAN TAPES
(Temporary test for modules (1) and
(2) with previous CLEANUP output)
(3)
CLEANUP I/O
ROUTINES FOR INCOS TAPES
(Useable on either 370 or UNIVAC)
≈ 200 Fortran Source Statements

Figure 22. Modules required for transfer of CLEANUP program to UNIVAC.

Rather than have CLEANUP call the I/O routines directly, it was decided to create an intermediate step consisting of an I/O interface called by CLEANUP which would in turn call the I/O routines. This was necessary because the only data used by CLEANUP for internal calculations are individual ion intensities while the INCOS tape data contained much additional identification information. Although not used by CLEANUP, these items were carried along to be able to write output spectra in the same format as on the original input tape. The interface served the dual function of keeping track of data not used directly inside CLEANUP and of making the I/O routines more versatile for use with other tasks.

The modification of the CLEANUP program involved roughly 500 statements: 100 statements for the rewritten routines inside CLEANUP, 200 statements for the I/O interface, and 200 statements for the I/O routines. To assure accuracy and quality control, several cross checks were made.

In order to insure the correctness of the modified CLEANUP source (1) and the I/O interface (2) a temporary set of I/O routines for reading the VARIAN CH 7 data (4) were written. These routines purposefully had exactly the same calling arguments as the INCOS tape I/O routines (3). These routines were written in standard FORTRAN and involved little more than READ and WRITE statements. It was then possible to test CLEANUP and the I/O interface on one of the eight trial runs. The prototype run was much more difficult to achieve than anticipated, mainly due to communication problems between the main CLEANUP source module and the I/O interface. However, because everything was in strict FORTRAN it was possible to use the WATFIV interpreter to find the majority of the CLEANUP bugs. The WATFIV interpreter again proved itself invaluable, not only in solving otherwise very subtle problems but also by insuring machine independence of the code. Thus, output of the CLEANUP source with the new I/O interface was verified against earlier output from the version of the CLEANUP program which was "spliced" into the RTI system.

It remained to test out the I/O routines (3) on an INCOS tape data set. After we were convinced that the I/O routines were operating properly, a program was written to generate a TIC plot of the first INCOS test run with the aid of the UNC plotter facility and the TUCC plotter routines. A reproduction of this plot is shown in Figure 23. This plot confirms that the data were correctly read and interpreted by the program.

Initially, running CLEANUP with INCOS test data on the TUCC IBM/370 detected only 8 eluants in the entire run. Closer examination of the data revealed that the average peak width for the INCOS runs extended over some 25 scans. Since the CLEANUP program derives its model singlets from a 7 scan window, not enough peak profile was available to obtain a good model shape (Figure 24). The simplest solution to this problem was to incorporate another parameter into the CLEANUP program which would average over a specified number of input spectra before any internal calculations were made. Once this modification was incorporated into CLEANUP, the same run was processed again, averaging over three spectra at a time. The results were much improved, showing a total of 72 eluants detected in the run (Figure 25).

After verification that the scan numbers of detected eluants agreed with the TIC plot (Figure 23) and checking the masses and intensities of detected eluants on the output tape against CLEANUP diagnostic printout, the task of duplicating this run with the same input tape on the UNIVAC computer was undertaken. A card copy of the CLEANUP source code was generated on the TUCC 370 and transported to the EPA UNIVAC when the source was recompiled and cross checking of its output. The chemist's printout of the test run on the UNIVAC is shown in Figure 26. The small discrepancies between this printout and the chemist's printout generated at TUCC are of a round-off type arising from hardware differences (i.e. word length, floating point arithmetic) between the two machines. These differences are too small to affect the performance of CLEANUP.

A second problem of greater concern occurred when trying to process multiple runs on a single tape with the EPA UNIVAC. Problems were encountered either when attempting to write a file mark at the end of

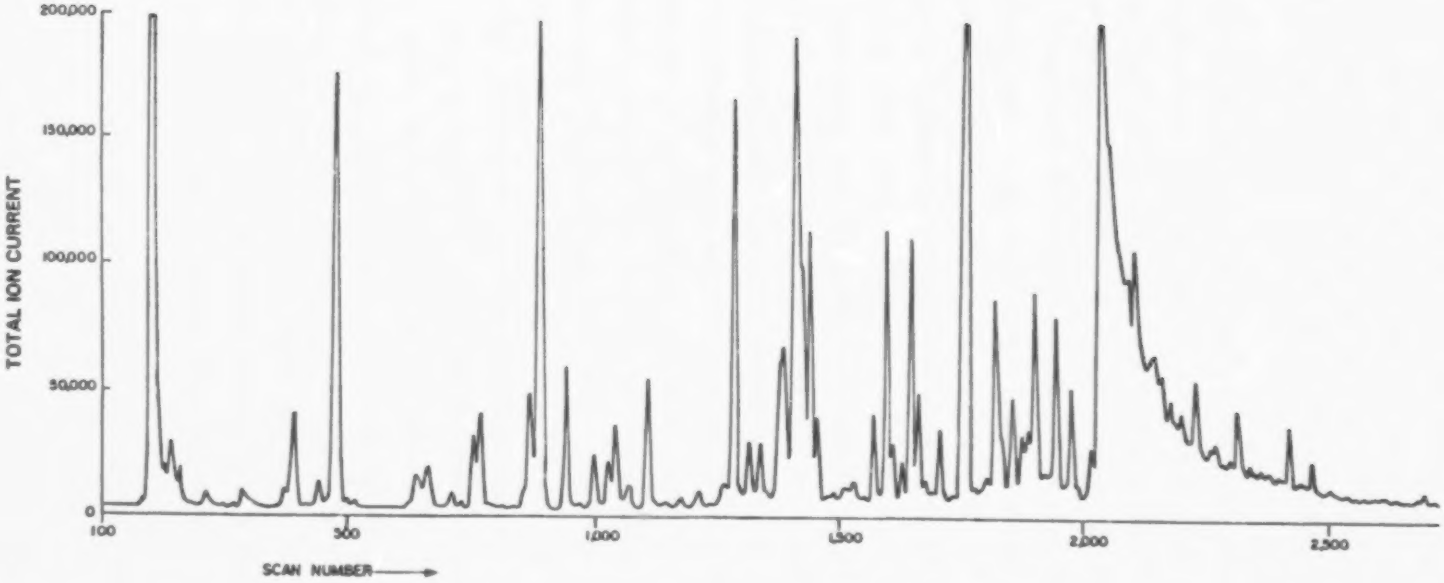
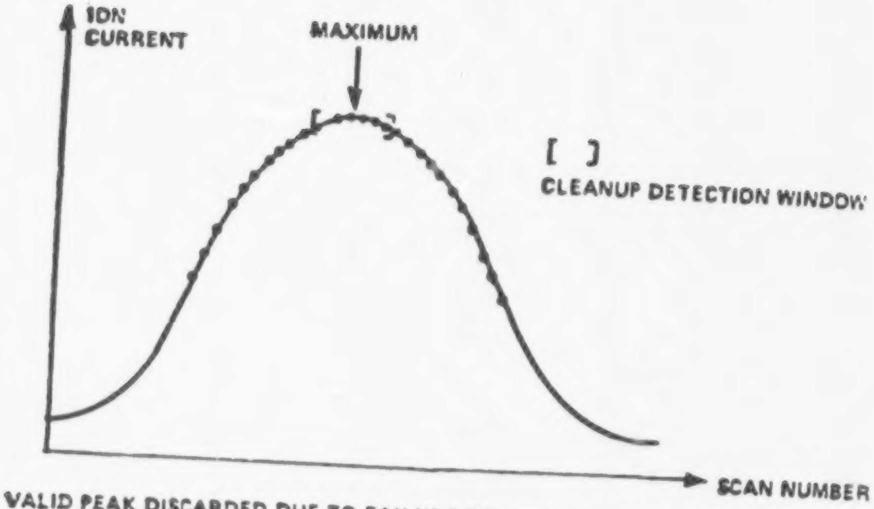
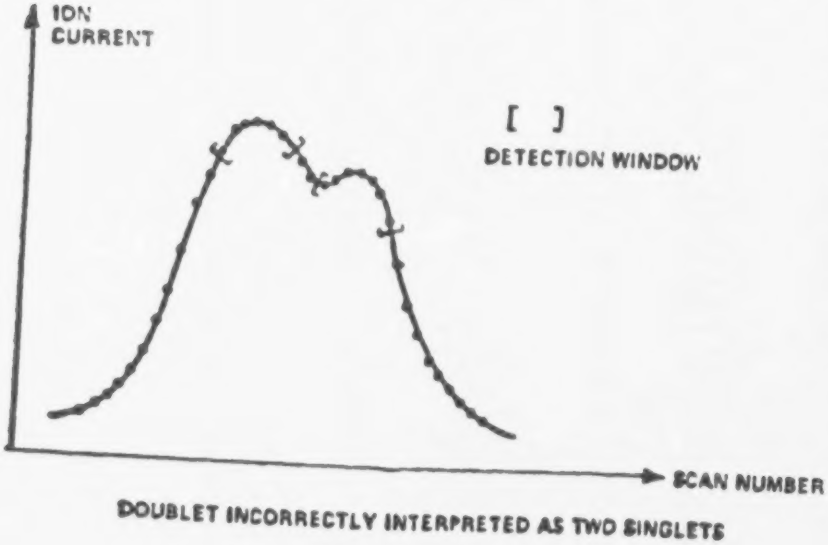


Figure 23. Raw TIC of INCOS test data.

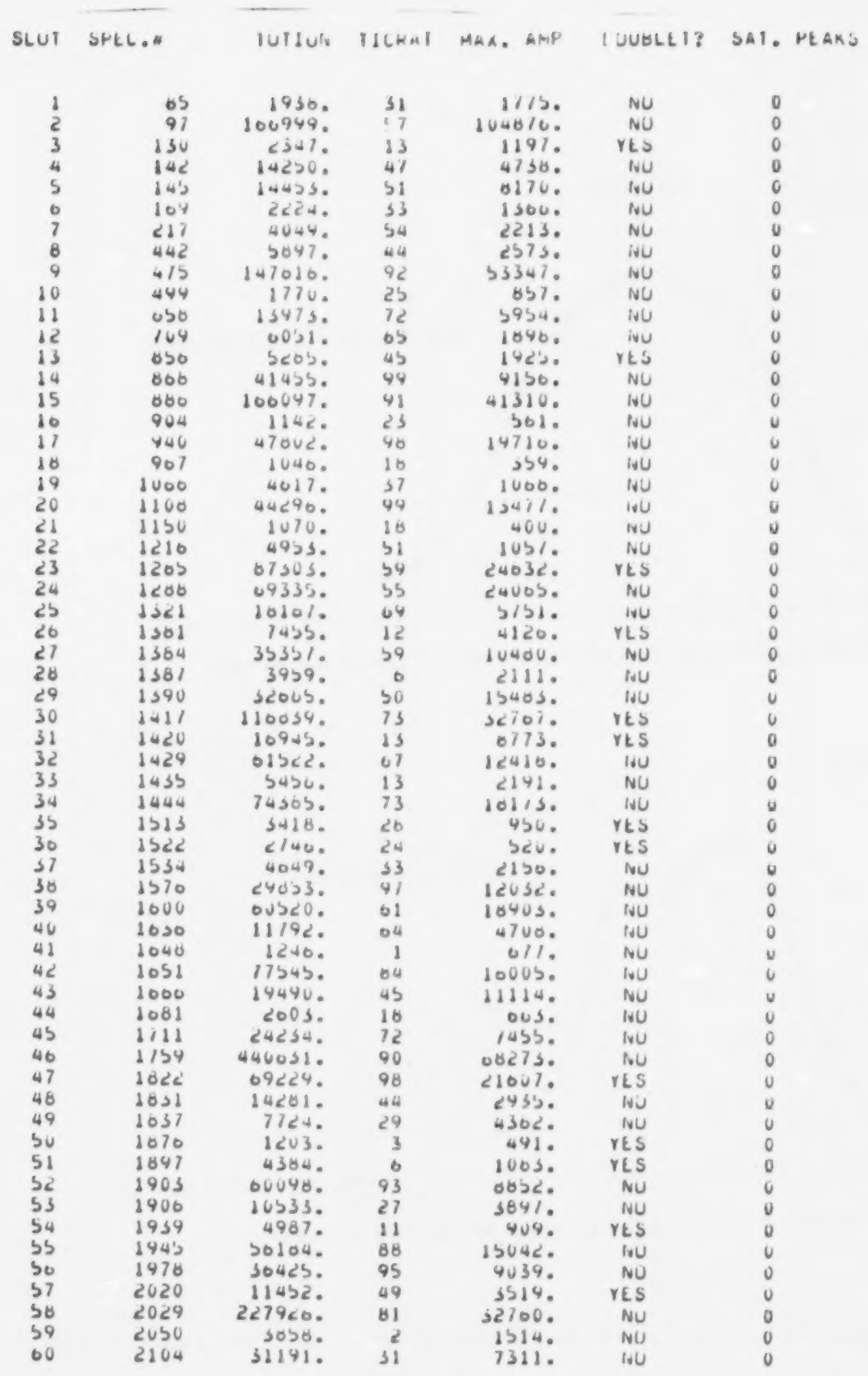


VALID PEAK DISCARDED DUE TO FAILURE OF LINEAR ERROR RATIO TEST



DOUBLET INCORRECTLY INTERPRETED AS TWO SINGLETS

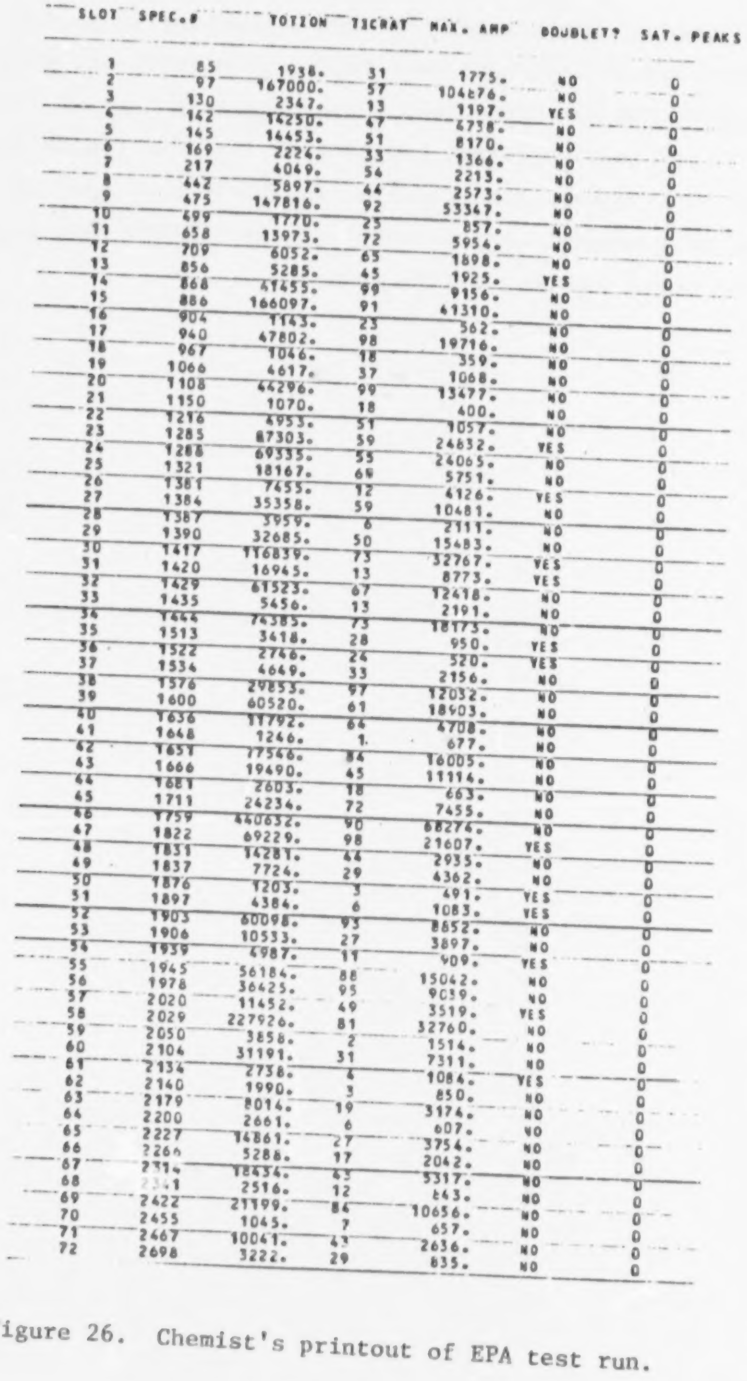
Figure 24. Possible consequences of scan rate too rapid for CLEANUP.



```
SLOT  SPEC.#      TOTION  TICRAT  MAX. AMP   DOUBLET?  SAT. PEAKS

   1      65       1936.    31      1775.      NO          0
   2      97     166949.   [illegible]7     104876.      NO          0
   3     130       2347.    13      1197.     YES          0
   4     142      14250.    47      4738.      NO          0
   5     145      14453.    51      6170.      NO          0
   6     169       2224.    33      1366.      NO          0
   7     217       4049.    54      2213.      NO          0
   8     442       5897.    44      2573.      NO          0
   9     475     147616.    92     53347.      NO          0
  10     499       1770.    25       857.      NO          0
  11     658      13973.    72      5954.      NO          0
  12     709       6051.    65      1896.      NO          0
  13     856       5265.    45      1925.     YES          0
  14     866      41455.    99      9156.      NO          0
  15     886     166097.    91     41310.      NO          0
  16     904       1142.    23       561.      NO          0
  17     940      47602.    98     14716.      NO          0
  18     967       1046.    16       359.      NO          0
  19    1066       4617.    37      1066.      NO          0
  20    1108      44296.    99     13477.      NO          0
  21    1150       1070.    18       400.      NO          0
  22    1216       4953.    51      1057.      NO          0
  23    1265      67303.    59     24632.     YES          0
  24    1288      69335.    55     24065.      NO          0
  25    1321      18167.    64      5751.      NO          0
  26    1361       7455.    12      4126.     YES          0
  27    1384      35357.    59     10480.      NO          0
  28    1387       3959.     6      2111.      NO          0
  29    1390      32065.    50     15483.      NO          0
  30    1417     116039.    73     32767.     YES          0
  31    1420      16945.    13      6773.     YES          0
  32    1429      61522.    67     12416.      NO          0
  33    1435       5456.    13      2191.      NO          0
  34    1444      74365.    73     16173.      NO          0
  35    1513       3418.    26       950.     YES          0
  36    1522       2746.    24       520.     YES          0
  37    1534       4649.    33      2156.      NO          0
  38    1576      24853.    97     12032.      NO          0
  39    1600      60520.    61     18903.      NO          0
  40    1636      11792.    64      4708.      NO          0
  41    1648       1246.     1       677.      NO          0
  42    1651      77545.    84     16005.      NO          0
  43    1666      19490.    45     11114.      NO          0
  44    1681       2603.    18       663.      NO          0
  45    1711      24234.    72      7455.      NO          0
  46    1759     446031.    90     68273.      NO          0
  47    1822      69229.    98     21607.     YES          0
  48    1831      14281.    44      2933.      NO          0
  49    1837       7724.    29      4362.      NO          0
  50    1876       1203.     3       491.     YES          0
  51    1897       4384.     6      1063.     YES          0
  52    1903      60098.    93      6852.      NO          0
  53    1906      10533.    27      3897.      NO          0
  54    1939       4987.    11       909.     YES          0
  55    1945      56164.    88     15042.      NO          0
  56    1978      36425.    95      9039.      NO          0
  57    2020      11452.    49      3519.     YES          0
  58    2029     227926.    81     32760.      NO          0
  59    2050       3658.     2      1514.      NO          0
  60    2104      31191.    31      7311.      NO          0
```

Figure 25. Chemist's printout of EPA test run generated at TUCC.



```
 SLOT  SPEC.#     TOTION  TICRAT  MAX. AMP   DOUBLET?  SAT. PEAKS

    1      85      1938.    31      1775.     NO          0
    2      97    167000.    57    104676.     NO          0
    3     130      2347.    13      1197.     YES         0
    4     142     14250.    47      4738.     NO          0
    5     145     14453.    51      8170.     NO          0
    6     169      2224.    33      1366.     NO          0
    7     217      4049.    54      2213.     NO          0
    8     442      5897.    44      2573.     NO          0
    9     475    147816.    92     53347.     NO          0
   10     499      1770.    25       857.     NO          0
   11     658     13973.    72      5954.     NO          0
   12     709      6052.    65      1898.     NO          0
   13     856      5285.    45      1925.     YES         0
   14     868     41455.    99      9156.     NO          0
   15     886    166097.    91     41310.     NO          0
   16     904      1143.    23       562.     NO          0
   17     940     47802.    98     19716.     NO          0
   18     967      1046.    18       359.     NO          0
   19    1066      4617.    37      1068.     NO          0
   20    1108     44296.    99     13477.     NO          0
   21    1150      1070.    18       400.     NO          0
   22    1216      4953.    51      1057.     NO          0
   23    1285     87303.    59     24632.     YES         0
   24    1288     69335.    55     24065.     NO          0
   25    1321     18167.    69      5751.     NO          0
   26    1381      7455.    12      4126.     YES         0
   27    1384     35358.    59     10481.     NO          0
   28    1387      3959.     6      2111.     NO          0
   29    1390     32685.    50     15483.     NO          0
   30    1417    116839.    73     32767.     YES         0
   31    1420     16945.    13      8773.     YES         0
   32    1429     61523.    67     12418.     NO          0
   33    1435      5456.    13      2191.     NO          0
   34    1444     74385.    73     18173.     NO          0
   35    1513      3418.    28       950.     YES         0
   36    1522      2746.    24       520.     YES         0
   37    1534      4649.    33      2156.     NO          0
   38    1576     29853.    97     12032.     NO          0
   39    1600     60520.    61     18903.     NO          0
   40    1636     11792.    64      4708.     NO          0
   41    1648      1246.     1.      677.     NO          0
   42    1651     77546.    84     16005.     NO          0
   43    1666     19490.    45     11114.     NO          0
   44    1681      2603.    18       663.     NO          0
   45    1711     24234.    72      7455.     NO          0
   46    1759    440632.    90     68274.     NO          0
   47    1822     69229.    98     21607.     YES         0
   48    1831     14281.    44      2935.     NO          0
   49    1837      7724.    29      4362.     NO          0
   50    1876      1203.     3       491.     YES         0
   51    1897      4384.     6      1083.     YES         0
   52    1903     60098.    93      8852.     NO          0
   53    1906     10533.    27      3897.     NO          0
   54    1939      4987.    11       909.     YES         0
   55    1945     56184.    88     15042.     NO          0
   56    1978     36425.    95      9039.     NO          0
   57    2020     11452.    49      3519.     YES         0
   58    2029    227926.    81     32760.     NO          0
   59    2050      3858.     2      1514.     NO          0
   60    2104     31191.    31      7311.     NO          0
   61    2134      2738.     4      1084.     YES         0
   62    2140      1990.     3       850.     NO          0
   63    2179      8014.    19      3174.     NO          0
   64    2200      2661.     6       607.     NO          0
   65    2227     14861.    27      3754.     NO          0
   66    2266      5288.    17      2042.     NO          0
   67    2314     18434.    43      5317.     NO          0
   68    2341      2516.    12       643.     NO          0
   69    2422     21199.    84     10656.     NO          0
   70    2455      1045.     7       657.     NO          0
   71    2467     10041.    43      2636.     NO          0
   72    2698      3222.    29       835.     NO          0
```

Figure 26. Chemist's printout of EPA test run.

the output cleaned up spectra tape or when attempting to locate the beginning of a subsequent run on the input test tape. Because it was possible to detect file marks separating the three runs on the EPA test tape at TUCC with no difficulty, it was concluded that the problem lay with the UNIVAC. Notwithstanding that the systems staff for the EPA UNIVAC have been alerted to these difficulties, we recommend that for the present CLEANUP runs on the UNIVAC be limited to one input and output file per tape.

SECTION 6

USER DESCRIPTION OF PROGRAM OPERATION

The final phase of implementing CLEANUP in the EPA environment consisted of testing the output tape written by the UNIVAC on the contracting agency's INCOS acquisition system. Although numerous problems were encountered, they were all of a systems oriented nature (incorrect tape density, incomplete format information, etc.) which commonly occur with machine transfer. No internal algorithm changes to the CLEANUP program were necessary in order to successfully read the test tape and obtain plausible output spectra with the INCOS software.

A user oriented description for running CLEANUP on the UNIVAC follows. To execute CLEANUP at the UNIVAC, it is necessary to deliver two tapes (one containing the INCOS raw input spectra and one for the output spectra) along with a card deck containing program control information (Figure 27). Cards 1 through 5 and 19 through 25 comprise the Executive Control Language necessary for the UNIVAC system to recognize and run the CLEANUP program. Of these, it is necessary to change the underlined portions of cards 19 and 20 with each run because they specify the tape serial numbers for CLEANUP to use. Cards 21 and 22 specify the positions of the data sets on the input and output tapes - in this example the input is read from the second file of tape B003KK and the output is written to the third file of tape B334KK. Cards 21 and 22 should be omitted if only one run is present on the input tape, and the cleaned up run is to be written on the beginning of the output tape. Card 1 - the "run" card contains the job name (CLNP) and the maximum time estimate in minutes (5). Although none of the present tests exceeded three minutes, it is conceivable that an increase of the time requirement could become necessary in the future.

```
@RUN 26CLNP/45/2,640265A010/KJK,5A010,5                           1
@PASSWD  KJK                                                      2
@ASG,A A.                                                         3
@ASG,T 4.       . CONTAINS THE CONTROL CARD DATA                  4
@DATA,I 4.      . THE ACTUAL CONTROL CARD DATA FOLLOWS            5
1     CLEANUP PRINT FLAG (IN I1)                                  6
0     DEBUG PRINT SWITCH IN I1 (0=NO,1=YES)                       7
 3    NUMBER OF INPUT SPECTRA TO AVERAGE (IN I2)                  8
6     MINIMUM THIRDS BETWEEN ELLUANTS (IN I1)                     9
   2    DETECTION WINDOW (IN I4)                                   10
  1000       MINIMUM FOR LESS THAN 5 PEAKS (IN I6)                11
  1000       MINIMUM FOR MORE THAN 4 PEAKS (IN I6)                12
   .0000     MINIMUM QUALIFYING RATE (IN F8.4)                    13
   .0350     ERROR RATIO (IN F8.4)                                14
  2          POINTS TO LEFT OF MODE (IN I3)                       15
   3         POINTS TO RIGHT OF MODE (IN I3)                      16
130000       SATURATION VALUE (IN F6.0)                           17
@END                                                              18
@ASG,TJN 7.,16D//////Q,B003KK        . INPUT TAPE                 19
@ASG,TJN/W 9.,16D//////Q,B334KK      . OUTPUT TAPE                20
@MOVE 7.,1 .SKIP 1 FILE ON INPUT TAPE-USE 2ND FILE                21*
@MOVE 9.,2 .SKIP 2 FILES ON OUTPUT TAPE-USE 3RD FILE              22*
@XQT,F A.CLEANUP                                                  23
@MARK 9.          . PUT A 'SAFETY' END OF FILE ON OUTPUT TAPE     24
@FIN             . END OF RUN                                     25
```

*omit until multiple run processing is implemented.

Figure 27. Example of CLEANUP control cards for EPA UNIVAC.

Cards 6 through 17 comprise the control cards used by the CLEANUP program. It is important to note that the variables shown here have been optimized with respect to the EPA test data taken on the INCOS system and hence differ from those in Table 7 which were optimized with respect to the VARIAN CH-7 at RTI. The most important difference is the additional control card (card 8) which instructs CLEANUP to average over 3 spectra. Also the instrument saturation value was changed to 130,000 (card 17), the minimum thirds between ellutants was increased from 3 to 6 (card 9), and the error ratio requirement was reduced from .0756 to .035 (card 14). The changes in minimum thirds and error ratio were made in order to accommodate the wider peak width and smoother data generated on the INCOS system.

SECTION 7

DELIVERY TO EPA

Accompanying this report are listings of the CLEANUP program which currently is being run on the UNIVAC, completing Phase IV of the contract. A backup tape of the same programs will be kept by RTI as well as the version of CLEANUP at TUCC which runs on INCOS tapes.

It still will be necessary for the contracting agency to carry out a series of test runs in order to determine the best operating parameters for CLEANUP. The system will respond differently to data derived from different mass spectrometers, as well as data run under differing GC conditions.

The program affords considerable flexibility in operation by the suitable adjustment of variable input parameters. These will still need to be exercised until optimum operational conditions are determined.

REFERENCES

1. "Mass Spectrometer--Computer System Particularly Suited for Gas-Chromatography of Complex Mixtures," R. A. Hites and K. Biemann, Anal. Chem., 40, 1217 (1968).

2. "Computer Techniques for Identifying Low Resolution Mass Spectra," S. L. Grotch, Anal. Chem., 43, 1362 (1971).

3. "Reconstructed Mass Spectra. A Novel Approach for the Utilization of Gas Chromatograph-Mass Spectrometer Data." J. E. Biller and K. Biemann, Anal. Lett., 7, 515 (1974).

4. "Extraction of Mass Spectra Free of Background and Neighboring Component Contributions from Gas Chromatography/Mass Spectrometry Data," R. G. Dromey, M. J. Stefik, T. C. Rindfleisch, and A. M. Duffield, Anal. Chem., 48, 1368 (1976).

5. Stephen M. Pizer, "Numerical Computing and Analysis," Science Research Associates, Inc., Chicago, 1975, p. 358ff.

APPENDIX A
CLEANUP SOURCE LISTING

```
      SUBROUTINE EPARD                                                  00035690
C                                                                       00035700
C   THIS ENTRY POINT READS ONE SCAN'S WORTH OF DATA PER CALL AND RETURNS00035710
C   IT TO THE CALLING ROUTINE IN COMMON BLOCKS IHDR AND INDATA.         00035720
C                                                                       00035730
      COMMON /HDR1/ NAME.HDRFLG.DATE.RUNHR.RUNMIN                       00035740
      COMMON /HDR2/ SAMPID.INAM                                         00035750
      COMMON /HDR3/ RUNCON.SECSCN                                       00035760
      COMMON /HDR4/ BUF4.LOPMAS.HIPMAS                                  00035770
      COMMON /IHDR/ JNAM.ISCN.IDBUF.IDHR.IDMIN.IRTMIN.IRTSEC.INMBAS.    00035780
     * IRBAS.IRTIC                                                      00035790
      COMMON /INDATA/ INBUF.INSEQ.NIN                                   00035800
      COMMON /OHDR/ ONAM.OSCN.ODBUF.ODHR.ODMIN.ORTMIN.ORTSEC.ONMBAS.    00035810
     * ORBAS.ORTIC                                                      00035820
      COMMON /OTDATA/ OTBUFF.OTSEQ.NOUT                                  0003583
      COMMON /FLAGS/ EOF                                                00035840
      COMMON /UNITS/ IRDR.IPTR1.IPTR2.IPRT3.IIN.IOUT                    00035850
      COMMON /IOBUF/ ITBUF1.CNTIN.CNTOUT.IOERR.ITBUF2                   00035860
      INTEGER*4 NAME(3).HDRFLG.DATE(2).RUNHR.RUNMIN.SAMPID(16).INAM(2). 00035870
     * RUNCON(16).BUF4(11).LOPMAS.HIPMAS.JNAM(3).ONAM(3).ISCN.OSCN.     00035880
     * IDBUF(2).ODBUF(2).IDHR.ODHR.IDMIN.ODMIN.IRTMIN.ORTMIN.IRTSEC.    00035890
     * ORTSEC.INMBAS.ONMBAS.EOF                                         00035900
     *.INSEQ.OTSEQ.NIN.NOUT.INBUF(411).OTBUFF(411)                      00035910
      INTEGER*4 RCDLEN/80/.CNTIN.CNTOUT.ITBUF1(20).IBLANK/4H    /.      00035920
     * INSPC(26).IDD(11).OTSPC(26).SCNEND/0/.ITBUF2(20)                 00035930
      INTEGER*4 OTSPCX(3.26).OTSEQX.OTSEQY
      REAL*4 SECSCN.ORBAS.IRBAS.IRTIC.ORTIC                             00035940
C                                                                       00035950
C   SET SCAN FLAG. COUNTER.                                             00035960
C                                                                       00035970
      SCNEND=0                                                          00035960
      NIN=0                                                             00035990
C                                                                       00036000
C   CLEAR TRANSFER ARRAY IN 'INDATA' BLOCK.                             00036010
C                                                                       00036020
      DO 5 I=1.411                                                      00036030
    5 INBUF(I)=0                                                        00036040
C                                                                       00036050
C   CALL BLKRD TO BRING A LOGICAL RECORD INTO THE BUFFER.               00036060
C                                                                       00036070
      CALL BLKRD                                                        00036080
      IF (IOERR.LT.0) GO TO 1000                                        00036090
C                                                                       00036100
C   DECODE A LOGICAL RECORD (ASSUMED TO BE A SCAN HEADER).              00036110
C                                                                       00036120
      DECODE(RCDLEN.5000.ITBUF1) JNAM.IDD(1).ISCN.IDD(2).IDBUF.IDHR.    00036190
     * IDD(3).IDMIN.IDD(4).IDD(5).IRTMIN.IDD(6).
     * IRTSEC.IDD(7).IDD(8).INMBAS.IDD(9).IRBAS.IDD(10).IDD(11).IRTIC   00036200
C                                                                       00036210
 5000 FORMAT(3A4.A1.I5.A2.2A4.I3.A1.I2.A4.A1.I3.A1.I2.A4.A2.I4.A1.F9.0. 00036220
     * A4.A1.F10.0)                                                     00036230
C                                                                       00036240
C   STOP IF NEW HEADER FOR RUN IS ENCOUNTERED.  (ISCN=0)                00036250
C                                                                       00036260
      IF (ISCN.EQ.0) GO TO 110                                          00036270
C                                                                       00036280
C   HANDLE SPECTRUM RECORD(S).  AT LEAST ONE IS ASSUMED TO BE PRESENT.  00036290
C                                                                       00036300
```

```
C
C   CHECK BUFFER VIA POINTERS TO DETERMINE AMOUNT OF REMAINING DATA.
C
C   RETURN HERE IF MORE THAN ONE SCAN RECORD.
C
  45 CALL BLKRD
     IF (IOERR.LT.0) GO TO 2000
C
C   DECODE A SPECTRUM RECORD.
C
  80 CONTINUE
C
C
     DECODE(RCDLEN.5005,ITBUF1) INSPC,INSEQ
C
5005 FORMAT(26I3,I2)
C
C   STORE THIS RECORD'S DATA IN INBUF USING SPECTRUM MASS AS INDEX OF IN
C
     DO 86 I=1.25,2
     J=I+1
     L=INSPC(I)
     IF (L.NE.0) GO TO 83
     SCNEND=1
     GO TO 86
  83 L=L-39
     IF (L.LT.1.OR.L.GT.411) GO TO 86
     INBUF(L)=INSPC(J)
     NIN=NIN+1
  86 CONTINUE
C
C   CHECK WHETHER END-OF-SCAN ENCOUNTERED (MASS VALUE = 0).
C
     IF (SCNEND.EQ.0) GO TO 45
C
C   RETURN WHEN SCAN ENDS (SCNEND = 1).
C
     SCNCNT=SCNCNT+1
     RETURN
C
C   ENTRY POINT FOR WRITING OUTPUT TAPE FOR INCOS USE.
C
     ENTRY EPAWRT
C
C   WRITE SCAN HEADER TO TEMPORARY BUFFER FROM COMMON BLOCK OHDR.
C
C
C
     ENCODE(RCDLEN.5000,ITBUF2) ONAM,IDD(1).OSCN,IDD(2).ODBUF,ODHR,
    * IDD(3).ODMIN,IDD(4),IDD(5).ORTMIN,IDD(6),
    * ORTSEC,IDD(7),IDD(8).ONMBAS,IDD(9).ORBAS,IDD(10),IDD(11).ORTIC
C
C
C
C   MOVE ENCODED HEADER RECORD TO OUTPUT BUFFER.  WRITE A FULL BUFFER AS
C   REQUIRED VIA NTRAN (CALLED FROM SUBROUTINE BLKWRT).
C
     CALL BLKWRT
C
C   STORE OTBUFF (PROCESSED OUTPUT SCAN) IN OTSPC AS MANY TIMES AS
```

```
C     WHEN OTSPC IS FULL. IT IS ENCODED INTO ITBUF2.
C
      J=-1
      OTSEQ=1
      DO 87 I=1.26
      OTSPC(I)=0
      DO 87 M=1.3
   87 OTSPCX(M.I)=0
      DO 100 I=1.411
C
C     LOOP ENDING AT 100 USES J & L AS INDICES TO OTSPC;
C
C     DATA IN OTSPC ARE:
C        MASS (I+39. WHERE 'I' IS THE INDEX OF OTBUFF);
C     AND INTENSITY. THE CONTENTS OF OTBUFF(I).
C
C     MASS IS STORED IN OTSPC(J) AND INTENSITY IS STORED IN OTBUFF(L).
C        WHERE L = J + 1.
C
      IF (OTBUFF(I).LE.0) GO TO 100
      J=J+2
      L=J+1
      OTSPC(J)=I+39
      OTSPC(L)=OTBUFF(I)
      IF (J.LT.25) GO TO 100
C
C     ENCODE SCAN DATA FROM OTSPC TO ITBUF2.
C
C
      DO 88 M=1.26
      I1=OTSPC(M)
      I2=I1/100
      I3=I1/10-I2*10
      I4=I1-I3*10-I2*100
      OTSPCX(1.M)=I2
      OTSPCX(2.M)=I3
   88 OTSPCX(3.M)=I4
      OTSEQX=OTSEQ/10
      OTSEQY=OTSEQ-OTSEQX*10
      ENCODE(RCDLEN.5040.ITBUF2) OTSPCX.OTSEQX.OTSEQY
 5040 FORMAT(80I1)
C
C
C     TRANSFER CARD IMAGE IN ITBUF2 TO OUTPUT BUFFER IN SUBROUTINE BLKWRT.
C
      CALL BLKWRT
C
C     CLEAR OTSPC AND INCREMENT OTSEQ.
C
      DO 90 K=1.26
      OTSPC(K)=0
      DO 90 M=1.3
   90 OTSPCX(M.K)=0
      OTSEQ=OTSEQ+1
      J=-1
  100 CONTINUE
C
C     SOME DATA REMAINS IN OTSPC (EVEN IF IT IS END-OF-SCAN ZEROS);
```

```
C   WRITE IT OUT.
C
      DO 103 M=1,26
      I1=OTSPC(M)
      I2=I1/100
      I3=I1/10-I2*10
      I4=I1-I3*10-I2*100
      OTSPCX(1,M)=I2
      OTSPCX(2,M)=I3
      OTSPCX(3,M)=I4
  103 CONTINUE
      OTSEQX=OTSEQ/10
      OTSEQY=OTSEQ-OTSEQX*10
      ENCODE(RCDLEN,5040,ITBUF2) OTSPCX,OTSEQX,OTSEQY
C
C   TRANSFER SCAN RECORD FROM ITBUF2 TO OUTPUT BUFFER.
C
      CALL BLKWRT
      RETURN
      ENTRY EPAHDR
C
C   ENTRY POINT TO READ FILE HEADER RECORDS (4) AND TRANSFER TO OUTPUT
C   BUFFER.
C
C HERE WE CHECK FOR A PREVIOUS EOF...
      IF(EOF.EQ.0)GO TO 105
      EOF=1
      RETURN
  105 CONTINUE
      CALL BLKRD
      IF (IOERR.LT.0) GO TO 2000
  110 CONTINUE
C
      DECODE(RCDLEN,5010,ITBUF1)        NAME,HDRFLG,DATE,RUNHR,RUNMIN
 5010 FORMAT(3A4,1X,I5,2X,2A4,I3,1X,I2,46X)
C
      DO 120 I=1,20
  120 ITBUF2(I)=ITBUF1(I)
      CALL BLKWRT
      CALL BLKRD
      IF (IOERR.LT.0) GO TO 2000
C
      DECODE(RCDLEN,5015,ITBUF1) SAMPID,INAM
 5015 FORMAT(16A4,10X,A4,A2)
C
      DO 130 I=1,20
  130 ITBUF2(I)=ITBUF1(I)
      CALL BLKWRT
      CALL BLKRD
      IF (IOERR.LT.0) GO TO 2000
C
      DECODE(RCDLEN,5020,ITBUF1)        RUNCON,SECSCN
 5020 FORMAT(16A4,10X,F6.2)
C
      DO 140 I=1,20
  140 ITBUF2(I)=ITBUF1(I)
      CALL BLKWRT
      CALL BLKRD
      IF (IOERR.LT.0) GO TO 1000
C
      DECODE(RCDLEN.5025.ITBUF1)       BUF4.LOPMAS.HIPMAS
 5025 FORMAT(6X.2A4.6X.2A4.6X.2A4.6X.5A4.5X.I3.1X.I3)
C
      DO 150 I=1.20
  150 ITBUF2(I)=ITBUF1(I)
      CALL BLKWRT
      RETURN
C
C   END-OF-FILE PROCESSING (IOERR < 0).
C
 1000 IF (IOERR.NE.-2) GO TO 2000
      EOF=1
      CALL BLKWRT
      CALL NTRAN$(IOUT.9.22)
      WRITE(IPTR1.5030) CNTIN.CNTOUT.EOF
 5030 FORMAT(' AT EOF CNTIN =',I6,'. CNTOUT =',I6,' EOF = ',I10)
      RETURN
C
C   ERROR PROCESSING.
C   IOERR<-2
C
 2000 EOF=-IOERR
      CALL BLKWRT
      CALL NTRAN$(IOUT.9.22)
      WRITE(IPTR1.5035) CNTIN. CNTOUT. EOF
 5035 FORMAT(' AT ERROR CNTIN =',I6,'. CNTOUT =',I6,' EOF = ',I10)
      RETURN
      END
      SUBROUTINE BLKRD
C
C   THIS SUBROUTINE TRANSFERS INPUT DATA FROM THE INCOS TAPE OF 120-WORD
C   (512-CHARACTER) BLOCKS INTO A 20-WORD (80-CHARACTER) BUFFER.
C
C   WHEN THE NTRAN BUFFER NO LONGER CONTAINS ENOUGH DATA FOR ONE LOGICAL
C   RECORD (<20 WORDS), THE TAIL OF THE BUFFER IS WRAPPED AROUND TO THE
C   BEGINNING OF THE BUFFER AND A NEW BLOCK READ INTO THE END OF THE
C   BUFFER VIA A CALL TO NTRAN$.
C
      COMMON /UNITS/ IRDR.IPTR1.IPTR2.IPRT3.IIN.IOUT
      COMMON /IOBUF/ ITBUF1.CNTIN.CNTOUT.IOERR.ITBUF2
      INTEGER*4 ITBUF1(20). CNTIN.CNTOUT.IOERR.ITBUF2(20)
      INTEGER*4 IISTRT/1/. BLKLEN/120/.RCDLEN/20/.IIEND/0/.
     * IBUF(140)
C
C   CHECK STARTING POSITION IN BUFFER.  IF IISTRT > (IIEND-RCDLEN+1)
C   READ A NEW BLOCK.
C
      IF (IISTRT.LE.(IIEND -RCDLEN+1)) GO TO 40
      IF (IISTRT.EQ.(IIEND+1)) GO TO 20
C
C   IF BUFFER HAS INSUFFICIENT DATA TO READ. MOVE RECORD TO BEGINNING OF
C   BUFFER AND STRING A NEW INPUT BLOCK BEHIND IT.
C
      J=0
      DO 10 I=IISTRT.IIEND
      J=J+1
```

```
   10 IBUF(J)=IBUF(I)
      IISTRT=J+1
      GO TO 30
   20 IISTRT=1
C
C     CALL BLOCK READ SUBROUTINE (SIMULATED FOR IBM 370 TESTING).
C
   30 CALL NTRAN$(IIN.3.BLKLEN.IBUF(IISTRT).IOERR.22)
      CNTIN=CNTIN+1
C
C     UPDATE BUFFER LENGTH TO REFLECT NEWLY READ DATA.
C
      IIEND =IISTRT+BLKLEN-1
      IISTRT=1
C
C     PROCESS NEXT LOGICAL RECORD.
C
   40 J=0
      IEND=IISTRT+19
      DO 50 I=IISTRT.IEND
      J=J+1
      ITBUF1(J)=IBUF(I)
   50 CONTINUE
C     CALL TRACE(IISTRT)
      IISTRT=IEND+1
      RETURN
      END
      SUBROUTINE BLKWRT
C
C     THIS SUBROUTINE FILLS A 512-CHARACTER (128-WORD) OUTPUT BUFFER
C     WITH 80-CHARACTER RECORDS.  WRAPAROUND IS USED TO MAKE THE BUFFER
C     COMPLETELY FULL AS REQUIRED BY INCOS.  THE BLOCKS, WHEN FULL, ARE
C     ACTUALLY WRITTEN ON TAPE BY SUBROUTINE NTRAN$.
C
      COMMON /UNITS/ IRDR.IPTR1.IPTR2.IPRT3.IIN.IOUT
      COMMON /FLAGS/EOF
      COMMON /IOBUF/ ITBUF1.CNTIN.CNTOUT.IOERR.ITBUF2
      INTEGER*4 ITBUF1(20). CNTIN.CNTOUT.IOERR.ITBUF2(20)
      INTEGER*4 IBUF(20).       OBUF(148)/148*4H    /.BLKLEN/128/.
     * RCDLEN/80/.IOSTRT/1/.IBLANK/4H    /.IOEND/20/
      EQUIVALENCE (IBUF(1).ITBUF2(1))
C
C     CHECK TO SEE WHETHER BUFFER IS ALREADY FULL TO 512 CHARACTERS.
C
C     ALSO. PUT OUT THE LAST BLOCK(S) (THERE MAY BE TWO REQUIRED) AT END-
C     OF-FILE OR ERROR IN READING OR WRITING.
      IF (IOSTRT.GT.128.OR.EOF.NE.0) GO TO 30
C
C     STORE FILLED TRANSFER BUFFER INTO OUTPUT BUFFER OBUF.
C
   10 IOEND=IOSTRT+19
      J=0
      DO 20 I=IOSTRT,IOEND
      J=J+1
   20 OBUF(I)=IBUF(J)
      IOSTRT=IOEND+1
      RETURN
C
```

```
C
C   UPON ENTERING AND FINDING A FULL BUFFER. WRITE IT TO TAPE.
C                                                                       00039270
   30 CALL NTRAN$(IOUT.1.BLKLEN.OBUF(1).IOERR.22)                       00039280
   31 CNTOUT=CNTOUT+1                                                   00039290
C                                                                       00039310
C   ADJUST CONTENTS OF OBUF SO THAT IT CONTAINS THE UNWRITTEN PORTION OF00039320
C   A 80-CHARACTER RECORD.  FILL THE REST OF OBUF WITH BLANKS.          00039330
C                                                                       00039340
   32 IOSTRT=1                                                          00039350
   33 IF (IOEND.LE.128) GO TO 50                                        00039360
   34 J=0                                                               00039370
      DO 40 I=129.IOEND                                                 00039380
      J=J+1                                                             00039390
   40 OBUF(J)=OBUF(I)                                                   00039400
      IOSTRT=J+1                                                        00039410
      IOEND=J                                                           00039420
C                                                                       00039430
C   STORE BLANKS IN REMAINDER OF OBUF.                                  00039440
C                                                                       00039450
   50 DO 60 I=IOSTRT.148                                                00039460
   60 OBUF(I)=IBLANK                                                    00039470
C                                                                       00039480
C   GET THE LAST BIT OF DATA ONTO TAPE IF EOF OR ERROR.                 00039510
C                                                                       00039520
      IF (IOEND.LT.IOSTRT.AND.EOF.GT.0) GO TO 30                        00039530
      IF (EOF.EQ.0) GO TO 10                                            00039490
C                                                                       00039500
C   OTHERWISE. FINISHED.                                                00039540
C                                                                       00039550
      RETURN                                                            00039560
      END                                                               00039570
      BLOCK DATA
      COMMON /FLAGS/ EOF                                                00039770
      INTEGER*4 EOF/0/                                                  00039780
      COMMON /IOBUF/ ITBUF1.CNTIN.CNTOUT.IOERR.ITBUF2                   00039790
      INTEGER*4 ITBUF1(20).CNTIN.CNTOUT.IOERR.ITBUF2(20)                00039800
      DATA CNTIN/0/.CNTOUT/0/.ITBUF1/20*4H    /.IOERR/0/.               00039810
     1ITBUF2/20*4H    /                                                 00039820
      END                                                               00039830
C THIS SUBROUTINE WAS CALLED ONLY FROM CUEXEC SO                        00039840
C CUEXEC WAS ELIMINATED AND ITS STATEMENTS WERE                         00000010
C INSERTED IN GCMSOL WHICH WAS IN TURN MADE THE MAIN PGM.               00000020
C                                          WM.F. HARGROVE 9/30/77       00000030
C..    SUBROUTINE GCMSOL(CLNFLG)                                        00000040
       INTEGER CLNFLG                                                   00000050
C                                                                       00000060
C       THIS COMPUTER PROGRAM WAS DEVELOPED WITH THE FUNDING SUPPORT     00000070
C       FROM THE NATIONAL INSTITUTES OF HEALTH (GRANTS RR-612 AND       00000080
C       GM-20832) AND THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION 00000090
       (GRANT NGR-05-020-632).                                          00000100
      INTEGER MASBUF(6).VALBUF(6).AJACEN.AHEAD.ISO(1000)                00000110
      INTEGER NAME.RDATE.BITMAP.GCTIME.LIBMAP.INTEG.LOMASS.MAD.ADS      00000120
      INTEGER HIMASS.POINTS.TAPCNT.MAXSP.LSTCNT.DEVICE.SPCNO            00000130
      INTEGER LSTMIN,SPEC.IND.IR.BK.NM.NTM.PNAME(32)                    00000140
      INTEGER HIST(10).UNIT.NNAME(18).IRNO                              00000150
      INTEGER NUSPEC.NMASS.OVLAP.EXPNAM.SRLIB(5).SECOND                 00000160
      INTEGER ODUBLT.OSATS.NLINE.RAWBUF.FSTPGE.FSTREC                   00000170
      INTEGER SATLST.SATINX.PASS1.ITOM.EOF                              00000180
      INTEGER GPEAK.SPARE1.SPARE2.SPARE3.SPARE4.SPARE7                  00000190
      INTEGER*4 TIMREM.TIMCNT
      REAL    SPARE5.SPARE6
      REAL    ISVP8.ISVP9
      REAL   QNORM
      REAL*8    C4.C1.C2.SGX.SHX
      DIMENSION G(9).H(11).GNU(18).HNU(18)
      COMMON PKHIST(20).NPHIST(20)
      COMMON YM5.YM4.YM3.YM2.YM1.Y0.YP1.YP2.YP3.YP4.YP5
      COMMON GM4.GM3.GM2.GM1.G0.GP1.GP2.GP3.GP4
      COMMON SG.SG2.SIG.SI.SI2.AN.AA.BB.CC.DD
      COMMON NAME(32).RDATE(5).BITMAP(64).GCTIME
      COMMON LIBMAP(17).INTEG.LOMASS.HIMASS.POINTS.TAPCNT
      COMMON MAXSP.LSTCNT.DEVICE(5).SPCNO
      COMMON SPEC(14.411).IND(3.411).NUSPEC(411).NMASS
      COMMON RMAX(3).IR(3).BK(3).NM(3).NTM(3).GPEAK(3.11)
      COMMON SPARE1.SPARE2.SPARE3.SPARE4.SPARE5.SPARE6.SPARE7
      COMMON RATMIN.RATEMX.SIGY0.SIGNXT.SIGLST
      COMMON ISVN8.ISVN9.ISVP8.ISVP9.SIGMA7.SATVAL.SATMAX
      COMMON MODE1.MODE2.NLWCUT.NUPCUT.SATPKS(25).NSAT.EXPNAM(5)
      COMMON /CSATLS/SATLST(25).SATINX
      COMMON /CAHEA/AHEAD(5)
      COMMON /PRNTFG/IPFLAG
      COMMON /MODES/ SGMODE(3).SHMODE(3).GMODE.HMODE
      COMMON /MASSES/ ISGMAS(3).ISHMAS(3).IGMASS.IHMASS
      COMMON /ERRORS/RMASS.ERRMAX.NERR.SIGERR.SECOND(411)
      COMMON /IFF/IFLAG(3)
      COMMON /AJACNT/AJACEN(2)
      COMMON /BACKER/SUMG.SUMG2
      COMMON /DUBBER/C4.C1.C2.SGX.SHX
      COMMON /PARAMS/MINTIC.MINNY.NTHIRD.RATM
      COMMON /HIND/IH1.IH2.MSINGS(2.50).MINX.LINGS(50).LINX
      COMMON /DUBFLG/IDUB
      COMMON /CRATMX/RATMXX
      COMMON /LFTRIT/ILEFT.IRIGHT
      COMMON /DUBBUG/IDEBUG
      COMMON /MONTOR/MONSPC.MONMAS
      COMMON /SLOTNO/NSLOT
      COMMON /CHEMPO/QDUBLT.QNORM.QSATS.NLINE
      COMMON /CFSTPG/FSTPGE
      COMMON /CRAWBU/RAWBUF(411)
      COMMON /LSQ/PASS1
      COMMON /CFSTRE/FSTREC
      COMMON /TOMS/ITOM
      COMMON /IMAD/MAD
      COMMON /IADS/ADS
      COMMON /NA/PNAME
      COMMON /UNITS/ IRDR.IPTR1.IPTR2.IPTR3.IIN.IOUT
C       IH1.IH2.IFLAG1.IFLAG2 USED IN TLOC AND ITS BRANCHES.
C HERE CLEANUP HAS BEEN GIVEN THE CAPABILITY OF MULTIPLE RUN
C PROCESSING - AND THE I/O UNIT NUMBERS HAVE BEEN PUT IN A
C COMMON BLOCK NAMED /UNITS/
C
C                                                  WM F HARGROVE
                                                   11/29/78
      IRDR=4
      IPTR1=6
      IPTR2=6
      IPTR3=6
      IIN=7
      IOUT=9
      IRNO=0
      CALL CUPRNT(3.NPHIST.MASBUF.PKHIST.AVE.
     1ERRMAX.ERR2.MODE1.MODE1.MODE2)
C GET CONTROL PARAMETERS - PARDIR HAS BEEN REWRITTEN AS OF
C                                                 12/12/78 WFH - RTI
      CALL PARDIR
C OPEN PRIMARY INPUT AND OUTPUT FILES
      CALL OPNFIL
C START OF RUN PROCESSING LOOP
   35 IRNO=IRNO+1
      EOF=0
      MU=0
      IH1 = 1
      IH2 = 2
      IFLAG(1) = 0
      IFLAG(2) = 0
      TAPCNT = 0
      AJACEN(2) = 0
C THIS MOD INSERTED TO HELP WITH UNINTIALIZED G ARRAY
C DOWN INSIDE SATLSQ. .....
C                                    RTI WFH 11/21/78
      DO 17 I=1.9
   17 G(I)=0.0
      DO 10 I = 1.5
10    AHEAD(I) = 0
C THIS MOD INSERTED TO INITIALIZE THE ARRAY GPEAK
C INSERTED 11/14/78 - AT RTI BY WFH
      DO 11 I=1.3
      DO 11 J=1.11
   11 GPEAK(I.J)=0
C THIS MOD TO INITIALIZE ARRAY MSINGS...
C INSERTED 4/12/78 - AT RTI BY WFH
      DO 13 I=1.2
      DO 13 J=1.50
   13 MSINGS(I.J)=0
C THIS MOD ALSO TO INITIALIZE AN ARRAY LINGS...
C INSERTED 4/13/78 - AT RTI BY WFH
      DO 14 I=1.50
   14 LINGS(I)=0
C THESE MODS ARE ALSO TO INITIALIXE ISVN8.ISVN9.ISVP8.ISVP9...
C INSERTED 4/14//78 - AT RTI BY WFH
      ISVP8=0.0
      ISVP9=0.0
      ISVN8=0
      ISVN9=0
C YET ANOTHER MOD TO INITIALIZE THE HISTOGRAMS - NPHIST.PKHIST
C INSERTED 4/14/78 - AT RTI BY WFH
      DO 15 I=1.20
      PKHIST(I)=0.0
   15 NPHIST(I)=0
C STILL ANOTHER MOD TO INITIALIZE THINGS FOR TLOC - SPARE3 TO SPARE7
C INSERTED 4/17/78 BY WFH AT RTI
      SPARE3=0
      SPARE4=0
      SPARE5=0.0
      SPARE6=0.0
C ONCE AGAIN AN UNINITIALIZED VARIABLE - THIS TIME IN LKAHED
```

```
C MOD INSERTED 4/28/78 AT RTI BY WFH
      OVLAP=0
      DO 12 I = 1.25
12    SATLST(I) = 0
      SATINX = 1
C THIS MOD INSERTED TO INITIALIZE ISGMAS FOR A
C MOD DOWN INSIDE TLOC TO PREVENT LOOKING FOR
C ELUTANTS WHICH HAVE NO SUITABLE MODEL....
C                     RTI - BY WFH 5/24/78
      DO 16 I=1.3
      ISHMAS(I)=-1
   16 ISGMAS(I)=-1
C      NSLOT.NLINE USED IN KAPUT VIA OUTDAT.
      NSLOT = 0
      NLINE = 0
      FSTPGE = 1
      FSTREC = 1
C      ANOTHER RANDOM VOID .. EH  .. HUI NENG ...
      NMASS=411
      LSTCNT=0
C HERE THE CALL TO PARDIR HAS BEEN SHIFTED TO THE TOP
C AND CALLING ARGUMENTS TO SETBUG HAVE BEEN MODIFIED
C                                         WM F HARGROVE
C                                         11/29/78
C     CALL PARDIR(CLNFLG.UNIT.NNAME)
      CALL SETBUG(NNAME.EOF)
      IF(EOF.NE.0)GO TO 50
C LOOP CONTROL MODIFIED SO PROGRAM WILL TERMINATE INSIDE MAIN
C CSECT - WFH RTI 9/31/78
C     MAXSP=MAXSP-10
      NS=13
   40 NS=NS+1
C THE UPDATING OF SPCNO IS NOW DONE INSIDE TLOC
C WHERE THE SPECTRUM RECORDS ARE INPUT...
C THIS WAS DONE SO THAT SPCNO COULD KEEP UP WITH
C GAPS IN THE SCAN NUMBERS...
C                                   WFH - RTI 6/15/78
C     SPCNO=SPCNO+1
C ALSO AN EOF FLAG WAS ADDED TO ENSURE TERMINATION INSIDE THE
C MAIN ROUTINE.....
C                                   WFH - RTI 9/31/78
      CALL TLOC(NS.M1.M2.M3.IFOUND.EOF)
      IF(EOF.EQ.1)GO TO 30
C      IFOUND = 0 IF NO SPECTRUM DETECTED.
      IF(IFOUND.EQ.0) GO TO 40
      MU=MU+1
      IF(MU.LE.1000)GO TO 23
      CALL CUPRNT(4.NPHIST.MASBUF.PKHIST.AVE.ERRMAX.
     1ERR2.SPCNO.MODE2.MODE2)
C     WRITE(IPTR2.24) SPCNO
C  24 FORMAT('-','*** CLEANUP TERMINATING DUE TO TABLE OVERFLOW ***'.
C    1/.1X.'LAST SPECRUM NO. IS '.I5)
      STOP 12
   23 ISO(MU)=SPCNO
20    CALL LKAHED(NS.M2.M3.OVLAP.G.H)
C      OVLAP = 1 IF ELUTANT DETECTED AHEAD.
      IF(OVLAP.EQ.0) GO TO 25
      CALL OFFSET(G.H.GNU.HNU.MODE1.MODE2)
```

```
C HOPEFULLY THE LAST FIX TO CUPRNT IN ORDER TO MAKE
C WATFIV HAPPY... THE ONLY ARGUMENTS USED IN THIS CALL ARE
C          MODE1 AND MODE2
C                                         WFH.RTI 4/26/78
      CALL CUPRNT(1.NPHIST.MASBUF.PKHIST.AVE.ERRMAX.
     1ERR2.MODE1.MODE2.MODE2)
C       ABOVE PRINTS MODE1 AND MODE2
25    CONTINUE
      ISKP=0
      PASS1 = 0
      NDIS=1
C       THIS PASS FOR MASSES DETECTED TO LEFT OF DETECTED SPECTRUM
C       AND NOT RESOLVED
C       IT SHOULD BE NOTED THAT THE DOUBLET RESOLVER IS ONLY RUN ON
C       PASS 2.
      CALL GETHGT(NS-2.M1.ISKP.OVLAP.NDIS.G.H.GNU.HNU)
C       THIS PASS FOR MASSES DETECTED AT THE SPECTRUM
C                                                ---
      NDIS=2
      CALL GETHGT(NS-1.M2.ISKP.OVLAP.NDIS.G.H.GNU.HNU)
      ISKP=1
      NDIS=3
C       THIS PASS FOR MASSES DETECTED TO THE RIGHT OF THE SPECTRUM
      CALL GETHGT(NS.M3.ISKP.OVLAP.NDIS.G.H.GNU.HNU)
C       PRINT LEAST SQUARE ERROR STUFF IF DOUBLET FOUND
      IF(OVLAP .EQ. 0)  GO TO 550
      IF (IDEBUG .EQ. 0) GO TO 550
      AVE = SIGERR/NERR
      ERRMAX = ERRMAX/NERR
      IMASS = RMASS
C HOPEFULLY THE LAST FIX TO CUPRNT IN ORDER TO MAKE
C C WATFIV HAPPY... THE ONLY ARGUMENTS USED IN THIS CALL ARE
C       AVE AND ERRMAX
C                                           WFH.RTI 4/26/78
      CALL CUPRNT(2.NPHIST.MASBUF.PKHIST.AVE.ERRMAX.ERR2.
     1MODE1.MODE2.MODE2)
C                    *   *                     *     *
C      WRITE OUT Q*H(IHSPEC) STUFF
      CALL CUPRNT(5.NPHIST.MASBUF.PKHIST.AVE.ERRMAX.
     1ERR2.SPCNO.MODE2.MODE2)
C     WRITE(IPTR2.900)
C     WRITE(IPTR2.910)
      ICOUNT = 0
      DO 600 I = 1.6
      MASBUF(I) = 0
600   VALBUF(I) = 0
      DO 650 I = 1.411
      IMASS = I + 39 - MAD
      IF(SECOND(I) .EQ. 32767) GO TO 650
      ICOUNT = ICOUNT + 1
      IF (ICOUNT .LT. 7) GO TO 620
C     WRITE(IPTR2.920) (MASBUF(J).VALBUF(J). J = 1.6)
      CALL CUPRNT(6.MASBUF.VALBUF.PKHIST.AVE.ERRMAX.ERR2.
     1SPCNO.MODE2.MODE2)
      ICOUNT = 1
      DO 610 J = 1.6
      MASBUF(J) = 0
610   VALBUF(J) = 0
```

```
620    MASBUF(ICOUNT) = IMASS
       VALBUF(ICOUNT) = SECOND(I)
650    CONTINUE
       IF ( MASBUF(1) .EQ. C) GO TO 550
C      WRITE(IPTR2.920) (MASBUF(J).VALBUF(J).J=1.6)
       CALL CUPRNT(6.MASBUF.VALBUF.PKHIST.AVE.ERRMAX.ERR2.
      1SPCNO.MODE2.MODE2)
C       THE ABOVE EMPTIES THE BUFFER.
550    CALL OUTDAT(SRLIB.NNAME)
       GO TO 40
C THIS MOD INSERTED IN CONJUNCTION WITH THE EOF FLAG ADDITION TO
C TLOC
C                                          WFH - RTI 9/31/78
C THE END OF THE CLEANUP MAIN CSECT HAS BEEN MODIFIED
C TO HANDLE MULTIPLE RUN PROCESSING.
C                                            WM F HARGROVE
C                                             11/29/78
C END OF SPECTRUM RUN
   30 CONTINUE
      CALL CLSRUN(ISQ.MU)
      GO TO 35
C END OF FILE AND TERMINATION OF CLEANUP PROGRAM
   50 CALL CUPRNT(8.NNAME.VALBUF.PKHIST.AVE.ERRMAX.
     2ERR2.TIMCNT.MODE2.MODE2)
      CALL CLSFIL
      STOP
      END
      SUBROUTINE BACLSQ(PKHGT)
C
C      THIS COMPUTER PROGRAM WAS DEVELOPED WITH THE FUNDING SUPPORT
C      FROM THE NATIONAL INSTITUTES OF HEALTH (GRANTS RR-612 AND
C      GM-20832) AND THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
C      (GRANT NGR-05-020-632).
C
C
C      *** THIS ROUTINE USES A MODEL PEAK AND CURRENT PEAK
C      *** AND DOES A LEAST SQUARES TO COMPUTE THE
C      *** BACKGROUND AT PEAK CENTRE (ASSUMES LINEAR BACKGROUND).
C
      INTEGER PKHGT.PASS1
      COMMON PKHIST(20).NPHIST(20)
      COMMON YM5.Y(9).YP5
      COMMON G(9)
      COMMON /LSQ/PASS1
      COMMON /BACKER/SG.SG2
      REAL   P.C
C
C      *** SOLUTION OF THREE LINEAR EQUATIONS.
C
C
C      P * SG**2 + C  = SGY
C
C      P * SG    + C  = SY
C
C
C         WHICH RESULT FROM THE LEAST SQUARE FITTING OF THE
C         ERROR TERMS.
C
```

```
C            E(I) = P * G(I) + C - Y(I)       ,  I = 1,9
C
C     WE SOLVE FOR P AND SET PEAKHGT = P * GTOP.
C
C       WHERE GTOP IS THE INTERPOLATED VALUE AT THE SPECTRUMS HISTOGRAM
C       POSITION.
C
      IF( PASS1 .EQ. 1 ) GO TO 10
      PASS1 = 1
      SG = 0
      SG2 = 0
      DO 5 I = 1.9
      SG = SG + G(I)/9
5     SG2 = SG2 + G(I) * G(I)/9
10    CONTINUE
      SY = 0
      SYG = 0
      DO 20 I = 1.9
      SY = SY + Y(I)/9
20    SYG = SYG + Y(I) * G(I)/9
      P = ( SYG - SY * SG )/( SG2 - SG * SG )
      PKHGT = P * G(5)
C      C = SY - P * SG
      RETURN
      END
      SUBROUTINE BLEED(Y.RATIO)
C      LEAST SQUARE FITS A LINE TO DATA TO DETERMINE ITS LINEARITY.
      REAL AVY.AXY.AXX.X.B.M.RATIO
      INTEGER N.I
      REAL Y(7).ERROR.FIT(7)
      AVY = 0
      AXY = 0
      AXX = 0
      DO 200 I = 1.7
      X = I - 4
      AVY = AVY + Y(I)
      AXY = AXY + X * Y(I)
200   AXX = AXX + X*X
      AVY = AVY/7
      AXX = AXX/7
      AXY = AXY/7
      B = AVY
      M = AXY/AXX
C      Y = M X  + B
C
C THE ERROR RATIO CALCULATION HAS BEEN MODIFIED SUBSTANTIALLY
C BECAUSE IT WAS FOUND TO BE HAYWIRE IN THE PREVIOUS VERSION.
C THE MOD IS A JOINT EFFORT OF DR/SHAH/PC/WFH
C                                         ENTERED BY WFH-RTI 9/31/78
      IF(AVY.EQ.0)GO TO 400
      RMS=0.0
      DO 300 I=1,7
      FIT(I)=M*(I-4)+B
      ERROR=FIT(I)-Y(I)
300   RMS=RMS+ERROR*ERROR
      RMS=SQRT(RMS)/7.0
      RATIO=RMS/AVY
      RETURN
C
  400 RATIO=0.0
      RETURN
      END
C      THIS SUBROUTINE DOES ALL PRINTING
C      FOR CLEANUP. PLEASE USE IT TO ADD NEW PRINTING
C      IN ANY SUBROUTINE.
C
C      THE CALL STATEMENT ALLOWS FOR PASSING OF ARRAYS
C      REALS AND INTEGERS.
C
C
C      THIS COMPUTER PROGRAM WAS DEVELOPED WITH THE FUNDING SUPPORT
C      FROM THE NATIONAL INSTITUTES OF HEALTH (GRANTS RR-612 AND
C      GM-20832) AND THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
C      (GRANT NGR-05-020-632).
C
      SUBROUTINE CUPRNT(N.IAR.IAR2.AR.R1.R2.R3.I1.I2.I3)
      INTEGER N.I1.I2.I3.IONE.NINE
      REAL R1.R2.R3
C THE DIMMENSIONS OF IAR AND AR HAVE BEEN CHANGED TO SUIT WATFIV
C THEY WERE AT 9....MUCH OF THIS CODE SEEMS NONSENSE...
C                                         WFH RTI 26/4/78
      DIMENSION IAR(1).AR(1).IAR2(1)
      COMMON /DUBBUG/IFLAG
      COMMON /UNITS/ IRDR.IPTR1.IPTR2.IPTR3.IIN.IOUT
      DATA IONE.NINE/1.9/
C
C      IFLAG IS RESET IF NO PRINTING IS TO BE DONE/
C
      IF (IFLAG.EQ.0) RETURN
      GO TO (10.140.150.160.170.180.200.210).N
C
C      THE ABOVE IS A CASE STATEMENT FOR EXECUTING PRINT
C      OPTIONS.C
C
C      GCMSOL RPINTING OF MODE1.MODE2
   10 IF(IFLAG.EQ.0)RETURN
      WRITE(IPTR2.510) I1.I2
      RETURN
C
  140 IF(IFLAG.EQ.0)RETURN
      WRITE(IPTR2.1400) R1.R2
      RETURN
C      FORMAT STATEMENTS FOLLOW.
C
510   FORMAT(/1X.'    HISTOGRAM LOCATIONS OF THE TWO SPECTRA ARE:'/
     -8X.'G-POSITION = '.I2.2X.'H-POSITION = '.I4)
1400  FORMAT(/' LEAST SQUARES ERROR STUFF FOR DUBLET:'/
     =' AVE. ERR.            AVE RMS ERR'
     =./.1X.E15.8.4X.E15.8.//)
C THESE PRINT STATEMENTS HAVE BEEN ADDED AS PART OF THE
C GENERAL REVAMPING OF CUPRNT....      WFH RTI 9/31/78
C
  150 WRITE(IPTR2.101)
  101 FORMAT('-'.'START OF CONVERTED MASS SPEC CLEANUP PROGRAM')
      RETURN
```

```
C
  160 WRITE(IPTR2,24) I1
   24 FORMAT('-','*** CLEANUP TERMINATING DUE TO TABLE OVERFLOW ***',
     1/,1X,'LAST SPECTRUM NO. IS ',I5)
      RETURN
C
  170 IF(IFLAG.EQ.0)RETURN
      WRITE(IPTR2.900)
900   FORMAT(//,20X,' COMPUTED H(MODE) = 0',/)
      WRITE(IPTR2.910)
910   FORMAT('  MASS VALUE  MASS VALUE  MASS VALUE  MASS VALUE  MASS VAL
     *UE  MASS VALUE'/)
      RETURN
C
C
  180 IF(IFLAG.EQ.0)RETURN
      WRITE(IPTR2.920) (IAR(I),IAR2(I),I=1.6)
  920 FORMAT(6(3X.I3.1X.I5))
      RETURN
C CUPRNT STATEMENTS 7 AND 8 HAVE BEEN MODIFIED TO
C FIT IN WITH THE SCHEME OF MULTIPLE RUN PROCESSING.
C                                              WM F HARGROVE
C                                              11/29/78 RTI
  200 WRITE(IPTR2.11)    I2.(IAR(I).I=1.18)
   11 FORMAT(1X.'END OF PROCESSING FOR RUN #'.I2.
     2' RUN TITLE: '.18A4)
      RETURN
C
  210 WRITE(IPTR2.102)
  102 FORMAT('-'.'***END OF CLEANUP PROCESSING***')
      RETURN
      END
      SUBROUTINE DETEKT (PKHIST.NPHIST.DTFLAG.NLWCUT.NUPCUT
     &    .NL.NU.IDXL.IDXU.INDX)
C
C       THIS COMPUTER PROGRAM WAS DEVELOPED WITH THE FUNDING SUPPORT
C       FROM THE NATIONAL INSTITUTES OF HEALTH (GRANTS RR-612 AND
C       GM-20832) AND THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
C       (GRANT NGR-05-020-632).
C
      INTEGER PLIMIT.DTFLAG.CPLMIT.SVFLAG.CASE1
      REAL MX.MX2.INTSUM
      DIMENSION PKHIST(20).NPHIST(20)
      COMMON /PARAMS/MINTIC.MINNY.NTHIRD.RATM
C      MINNY IS USED FOR FEWER THAN 5 PEAKS
C      REQUIRE  TOTAL ION CURRENT FOR DETECTED SPECTRUM TO BE MINTIC
C      IN AND AROUND THE MAX IF 5 PEAKS LOCALLY.
      COMMON /TOMS/ITOM
      COMMON /CAHEA/AHEAD(5)
      INTEGER AHEAD.ITOM
C      ********************************************
C      WE SCAN THE HISTOGRAMIC WINDOW *NL.NU* FOR A MAXIMUM
      DTFLAG=0
      INDX=NL
      NUM=NPHIST(NL)
      MX=PKHIST(NL)
C      THE HISTOGRAMIC WINDOW *IDXL.IDXU* IS WHERE A MAXIMUM
C      WILL OCCUR IF A SPECTRUM IS ELUTING.
```

```
      NUM2=NPHIST(IDXL)
      MX2=PKHIST(IDXL)
      INDX2=IDXL
C           MX.NUM = MAX +NL.NU@  &   MX2.NUM2 = MAX +IDXL.IDXU@
      DO 10 I=NL.NU
      IF ( MX .GE. PKHIST(I) ) GO TO 9
      MX=PKHIST(I)
      NUM=NPHIST(I)
      INDX=I
9     CONTINUE
C      NOTE: +IDXL.IDXU@ IS  WINDOW FOR DETECTING ELUTANT NOW
      IF ( I .LE. IDXL  .OR.  I .GT. IDXU )  GO TO 10
      IF ( MX2  .GE.  PKHIST(I) ) GO TO 10
C      GET MAXIMUM IN WINDOW  + IDXL.IDXU @
      MX2=PKHIST(I)
      NUM2=NPHIST(I)
      INDX2=I
10    CONTINUE
C      SEE IF IN RIGHT HISTOGRAM WINDOW. +IDXL.IDXU@
      IF ( INDX .GE. IDXL  .AND.  INDX .LE. IDXU ) GO TO 15
C      IF WE GET HERE THE ABSOLUTE MAXIMUM WAS NOT IN +IDXL.IDXU@
C      AND  IDXL  .LE.  INDX2   .LE. IDXU
C      REQUIRE: 1. THAT TIC (INDX2) BE LARGER THAN ITS IMMEDIATE
C                  NEIGHBORS.
C                2. THAT PKHIST(IDXU) + PKHIST(IDXU+1)
C                     + PKHIST(IDXL)     .GE.    MINTIC.
      IF ( MX2  .LT.  PKHIST(INDX2-1) )  RETURN
      IF ( MX2  .LT.  PKHIST(INDX2+1) ) RETURN
C      CHECK ON HOW CLOSE WE WILL ALLOW HISTOGRAM BUMPS.
C      DEFAULTS TO 2 SLOTS APART. SPECIAL FLAG ITOM IS
C      USED FOR FRAGMENTOGRAPHY RUNS TO TRY TO SEPARATE
C      DEUTERATED PEAKS FROM NON-DEUTERATED PEAKS.
C
C THIS IS A TEMPORARY MOD IN ORDER TO GATHER STATISTICS ON
C THE INPUT PARAMETER ITOM - CODE 6
C                                              UFH RTI 9/31/78
C
C     ISPC=0
C     ITYP=6
C     WRITE(20.6251) ITYP.ISPC
C6251 FORMAT(A4.A4)
C
      IF ( ITOM .EQ. 1 ) GO TO 79
      IF ( MX2  .LT.  PKHIST(INDX2-2) )  RETURN
      IF ( MX2  .LT.  PKHIST(INDX2+2) ) RETURN
C
C TEMPORARY MOD IN ORDER TO GATHER STATISTICS ON
C THE INPUT PARAMETER ITOM - CODE 7
C                                         UFH RTI 9/31/78
C     ISPC=0
C     ITYP=7
C     WRITE(20.6254) ITYP.ISPC
C6254 FORMAT(A4.A4)
79    CONTINUE
      PKSUM = PKHIST(IDXL) + PKHIST(IDXL+1) + PKHIST(IDXU)
      IF ( IDXL .NE. 4 ) GO TO 12
      IF ( AHEAD(1) .EQ. 1 ) PKSUM = MINTIC
C      THE ABOVE GUARANTEES DETECTION IF SPECTRUM WAS DETECTED
```

```
C       AHEAD.
12     CONTINUE
C
C THIS IS A TEMPORARY MOD IN ORDER TO GATHER STATISTICS ON
C THE INPUT PARAMETER ITOM - CODE 5
C                                                 WFH RTI 9/31/78
C      ISPC=0
C      ITYP=5
C      WRITE(20,6252) ITYP,ISPC,PKSUM
C6252 FORMAT(A4,A4,36X,A4)
       IF ( PKSUM .LT. MINTIC ) RETURN
       NLWCUT=INDX2-2
       NUPCUT=INDX2+2
       INDX=INDX2
C       DTFLAG = 1 SAYS WE'VE FOUND ONE
       DTFLAG=1
       RETURN
15     CONTINUE
C       WE NOW SUM NUMBER OF PEAKS AND TIC.
       IF ( IDXL .LT. 7 ) GO TO 16
C       MAKE SURE WE HAVE A HISTOGRAMIC BUBBLE
C        IF LOOKING AHEAD.
       BUBBLE = PKHIST(INDX)
       IF ( BUBBLE .LT. PKHIST(INDX-1) ) RETURN
       IF ( BUBBLE .LT. PKHIST(INDX-2) ) RETURN
16     CONTINUE
       PKSUM = NPHIST(IDXL)
       INTSUM = PKHIST(IDXL)
C       IF IDXL = 9 THEN WE ARE LOOKING AHEAD, AND AN EARLIER
C       ELUTANT WAS DETECTED AT 6. WE ONLY SCANNED +9,100.
       IF (IDXL .EQ. 9) GO TO 150
       PKSUM = PKSUM + NPHIST(IDXL + 1)
       INTSUM = INTSUM + PKHIST(IDXL + 1)
       IF ( IDXL .EQ. 8 ) GO TO 100
C       HERE WE ONLY SCANNED +8,100
       PKSUM = PKSUM + NPHIST(IDXU)
       INTSUM = INTSUM + PKHIST(IDXU)
       IF ( IDXL .EQ. 7 ) GO TO 100
C       HERE WE SCANNED +7,100. IN ALL SPECIAL CASES A ELUTANT WAS
C       DETECTED IN +4,5,60 SO WE DO NO LEFT BOUNDARY CHECKING.
C       SEE IF MAXIMUM OCCURRED ON LEFT BOUNDARY
       IF ( INDX .NE. IDXL ) GO TO 100
       PKSUM = PKSUM + NPHIST(IDXL-1)
       INTSUM = INTSUM + PKHIST(IDXL-1)
       GO TO 150
C       SEE IF MAXIMUM OCCURRED ON RIGHT BOUNDARY
100    CONTINUE
       IF ( INDX .NE. IDXU ) GO TO 150
       PKSUM = PKSUM + NPHIST(IDXU+1)
       INTSUM = INTSUM + PKHIST(IDXU+1)
150    CONTINUE
C       GUARANTEE SPECTRUM DETECTION IF DETECTED AHEAD.
       IF ( IDXL .NE. 4  ) GO TO 155
       IF ( AHEAD(1) .EQ. 0 ) GO TO 155
       PKSUM = MINTIC
       IF ( PKSUM .LT. MINNY ) PKSUM = MINNY
155    CONTINUE
C       CHECK FOR FEWER THAN 5 PEAKS
```

```
C
C THIS IS A TEMPORARY MOD IN ORDER TO GATHER STATISTICS ON
C THE INPUT PARAMETER ITOM - CODE 4
C                                          WFH RTI 9/31/78
C     ISPC=0
C     ITYP=4
C     WRITE(20.6257) ITYP.ISPC.PKSUM.INTSUM
C6257 FORMAT(A4.A4.28X.A4.A4)
      IF (PKSUM .GE. 5) GO TO 200
      IF( INTSUM .GE. MINNY ) GO TO 18
      RETURN
200   CONTINUE
C      IF AT LEAST 5 PEAKS. THEN REQUIRE INTSUM > MINTIC.
      IF(INTSUM .LT. MINTIC) RETURN
18    CONTINUE
      DTFLAG=1
      IF ( NL .NE. 1 ) RETURN
C      NL = 1  --> WE ARE NOT LOOKING AHEAD. AND NEED HISTOGRAMIC
C      BOUNDS. NLWCUT AND NUPCUT.
      SLEFT=0.0
      SRIGHT=0.0
      NUM1=INDX+1
      NUM3=INDX+3
C      GET HISTOGRAMIC "WEIGHTS" TO THE LEFT AND RIGHT OF THE
C      MAXIMUM AT PKHIST(INDX).
      J = 0
      DO 20 I=NUM1.NUM3
      J = J + 1
      JJ = INDX - J
      SLEFT = SLEFT + PKHIST(JJ)
      SRIGHT = SRIGHT + PKHIST(I)
20    CONTINUE
      NLWCUT = INDX - 2
      NUPCUT = INDX + 2
      IF ( SLEFT - SRIGHT ) 30.40.50
C      MORE "WEIGHT" TO RIGHT IF WE GET HERE
30    NUPCUT = NUPCUT + 1
      GO TO 40
C      MORE "WEIGHT" TO LEFT IF WE GET HERE
50    NLWCUT = NLWCUT - 1
40    RETURN
      END
      SUBROUTINE DUBLET (G.H.Y.STORE.M.IHSPEC.GHGT.ISWD.NS)
      REAL  GHGT
      INTEGER STORE.M.IHSPEC.ISWD.SECOND.NS.MAD
      REAL KK .LL.MM.NN
      COMMON /DUBBER/C4.C1.C2.SG.SH
      COMMON /MONTOR/MONSPC.MONMAS
      COMMON /ERRORS/RMASS.ERRMAX.NERR.SIGERR.SECOND(411)
      COMMON /IMAD/MAD
      COMMON /UNITS/ IRDR.IPTR1.IPTR2.IPTR3.IIN.IOUT
      INTEGER  132766
      REAL*3   C4.C1.C2.SG.SH
      REAL*8   SY.SGY.SHY.SG2.SH2.SGH
      REAL     ERR.ERRM
C      *** ROUTINE TO RESOLVE A DUBLET USING TWO MODEL
C      *** PEAKS BY A LEAST SQUARES METHOD. THE MODEL
C      *** INCLUDES A LINEAR BACKGROUND.
```

```
C        THE LEAST SQUARE NORMAL EQUATIONS ARE DERIVED FROM
C        THE FOLLOWING EQUATION FOR ERROR MINIMAZATION:
C           E(I) = P * G(I) + Q * H ( I ) + C - Y(I)  .I = 1.14
C             <X> = AVERAGE VALUE OF X
C        THIS LEADS TO THE FOLLOWING SYSTEM OF EQUATIONS :
C           P * <SG**2> + Q * <SGH>    + C * <SG> = <SGY>
C           P * <SGH>   + Q * <SH**2> + C * <SH> = <SHY>
C           P * <SG>    + Q * <SH>    + C        = <SY>
       REAL   G(18).H(18).Y(18)
       REAL   ERRT1(14).FIT(14)
       REAL   P.Q.C.ZERO.HMD(14).GMD(14)
       INTEGER ILOW.IUP
       DATA ZERO/0.0/
       DATA I32766/32760/
C       ***********************************HI THERE********************
C       INITIALIZE SOME VARIABLES
       ILOW = 3
       IUP = 11
C       WE TRUNCATE THE SECOND MODEL AT THE 11TH POSITION SO AS
C       ONLY TO USE POINT WHERE REAL DATA WAS SUPPLIED.
       AN = IUP - ILOW + 1
C         TEST TO SEE IF WE'VE ALREADY BEEN BY FOR THIS SPECTRUM
       IF( ISWD . EQ.  0) GO TO 20
       ISWD = 0
       SG2 = 0
       SGH = 0
       SG  = 0
       SH2 = 0
       SH  = 0
C       DO SOME              SUMS.
       DO 10 I = ILOW.IUP
       SG = SG + G(I)
       SG2 = SG2 + G(I) * G(I)
       SH = SH + H(I)
       SH2 = SH2 + H(I) * H(I)
       SGH = SGH + G(I) * H(I)
10     CONTINUE
C       MAKE ALL AVERAGE VALUES
       SG = SG/AN
       SG2 = SG2/AN
       SH = SH/AN
       SH2 = SH2/AN
       SGH = SGH/AN
C       NOW DO WHAT COMPUTING WE CAN FOR MODEL STUFF
       C4  = SG2 - SG * SG
       C1  = SGH - SG * SH
       C2  = SH2 - SH * SH
20     CONTINUE
       SY=0
       SGY=0
       SHY=0
       DO 30 I = ILOW.IUP
       SY=SY + Y(I)
       SGY=SGY + G(I) * Y(I)
       SHY=SHY + H(I) * Y(I)
30     CONTINUE
       SY = SY/AN
       SGY = SGY/AN
```

```
      SHY = SHY/AN
C      NOW SOLVE FOR P.Q.C
      C5  = SGY - SY * SG
      C3  = SHY - SY * SH
      P = ( C5 * C2 - C3 * C1 )/( C4 * C2 - C1 * C1 )
      Q = ( C3 - P * C1 )/C2
      C = SY - P * SG - Q * SH
      GHGT = P * G(7)
      HHGT = Q * H(IHSPEC)
      IF( GHGT .LT. ZERO ) GHGT = ZERO
      IF( HHGT .LT. ZERO ) HHGT = ZERO
C      DO MEAN SQUARE ERROR CALCULATION
      ERR = 0
      ERRM = 0
      DO 40 I = ILOW.IUP
      GMD(I)   = P * G(I)
      HMD(I)   = Q * H(I)
      FIT(I)   = GMD(I) + HMD(I) + C
      ERRT1(I) = Y(I) - FIT(I)
      ERRM     = ERRM + ERRT1(I)
40    ERR      = ERR  + ERRT1(I) * ERRT1(I)
      ERRM = ERRM/AN
      RMS = ERR/AN - ERRM * ERRM
      IF( RMS .LT. 0 ) RMS = 0
      ERR = SQRT(RMS)
      SIGERR = SIGERR + ERRM
      ERRMAX = ERRMAX + ERR
      NERR = NERR + 1
      STORE = 1
C      FLAG FOR SETGHY TO SIGNAL STORING OF NEW VALUE IN NUSPEC(M)
      IF(HHGT .GT. 132766) HHGT = 132766
      SECOND(M) = HHGT
200   CONTINUE
      RETURN
C THIS SECTION OF DUBLET DEALS WITH PRINTING
C OUT OF MONITORED SPECTRA AND MONITORED MASSES
C IT WAS COMMENTED OUT BECAUSE NO EASY WAY OF
C INCORPORATING THIS DIAGONISTIC INTO THE
C MULTIPLE RUN PROCESSING SCHEME COULD
C BE THOUGHT OF. IT MAY BE REINSTITUTED IN THE
C EVENT OF SUBTILE DOUBLET RESOLUTION PROBLEMS.
C                                     WM F HARGROVE
C                                     12/7/78 - RTI
C      IF ( MONMAS .EQ. 0 ) RETURN
C      IM = M + 39-MAD
C      IS = NS - 6
C      IF ( MONMAS .NE. IM   .OR. MONSPC .NE. IS) RETURN
C      WRITE(IPTR2.2500) ERR.ERRM
C      WRITE(IPTR2.2000) IM.P.Q.C
C      WRITE(IPTR2.2100) (Y(I).I = 3 . 11)
C      WRITE(IPTR2.2200) (G(I).I = 3 . 11)
C      WRITE(IPTR2.2300) (H(I).I = 3 . 11)
C      BAK = C
C      SQHT = GHGT + Q * H(7) + BAK
C      WRITE(IPTR2.2400) GHGT.BAK.SQHT
C999  RETURN
C2000 FORMAT(//.' DOUBLET STUFF FOR MASS = '.I3./
C    -.'      P             Q             C              './
C    -.3(2X.F10.3))
C2100 FORMAT(//.' Y-VALUES'./.5(2X.F7.0)./.4(2X.F7.0))
C2200 FORMAT(//.' G-VALUES'./.5(2X.F7.4)./.4(2X.F7.4))
C2300 FORMAT(//.' H-VALUES'./.5(2X.F7.4)./.4(2X.F7.4))
C2400 FORMAT(//.'    P * G(7)  BACK GROUND  P*G + Q*H + C '.
C    -/.3(4X.F8.0)/)
C2500 FORMAT(//.' RMS ERROR = '.E14.8.' MEAN ERROR = '.E20.8)
      END
      SUBROUTINE GETHGT(NS.MN.ISKP.OVLAP.NDIS.G.H.GNU.HNU)
      INTEGER PKHGT.MNB.MI.MN.ISKP.JMPK.NUHGT
      DIMENSION G(9).H(11).GNU(18).HNU(18)
C      *** THIS ROUTINE IS THE MAIN PROCESSING ROUTINE. IT
C      *** PICKS UP THE MASS INDICES THAT HAVE BEEN FLAGGED
C      *** IN THE DETECTION PHASE AND PROCEEDS TO  HANDLE
C      *** BACKGROUND REMOVAL AND DOUBLET RESOLUTION AND
C      *** THE PROCESSING OF SATURATED PEAKS. IT IS CALLED
C      *** THREE TIMES IN PROCESSING ANY ONE COMPONENT DUE TO
C      *** THE SPREAD THAT CAN EXIST IN PEAK MODES FOR ANY
C      *** COMPONENT (PARTICULARLY THE LOW INTENSITY ONES).
      INTEGER NAME.RDATE.BITMAP.GCTIME.LIBMAP.INTEG.LOMASS
      INTEGER HIMASS.POINTS.TAPCNT.MAXSP.LSTCNT.DEVICE.SPCNO
      INTEGER SPEC.IND.IR.BK.NM.NTM
      INTEGER GPEAK.SPARE1.SPARE2.SPARE3.SPARE4.SPARE7
      INTEGER ISWD.MAD
      INTEGER NUSPEC.NMASS.OVLAP
      INTEGER SECOND
      INTEGER PASS1
      INTEGER QDUBLT.QSATS.NLINE.BEDROK.AJACEN.SATLST.SATINX
      REAL     ISVP8.ISVP9
      REAL HIST(10).YVAL
      REAL     SPARE5.SPARE6
      REAL    QNORM.RATIO
      REAL    PEPCNT
      REAL*8 GTHT
      COMMON PKHIST(20).NPHIST(20)
      COMMON YVAL(11)
      COMMON GM4.GM3.GM2.GM1.G0.GP1.GP2.GP3.GP4
      COMMON SG.SG2.SIG.SI.SI2.AN.AA.BB.CC.DD
      COMMON NAME(32).RDATE(5).BITMAP(64).GCTIME
      COMMON LIBMAP(17).INTEG.LOMASS.HIMASS.POINTS.TAPCNT
      COMMON MAXSP.LSTCNT.DEVICE(5).SPCNO
      COMMON SPEC(14.411).IND(3.411).NUSPEC(411).NMASS
      COMMON RMAX(3).IR(3).BK(3).NM(3).NTM(3).GPEAK(3.11)
      COMMON SPARE1.SPAPE2.SPARE3.SPARE4.SPARE5.SPARE6.SPARE7
      COMMON RATMIN.RATEMX.SIGY0.SIGNXT.SIGLST
      COMMON ISVN8.ISVN9.ISVP8.ISVP9.SIGMA7.SATVAL.SATMAX
      COMMON MODE1.MODE2.NLWCUT.NUPCUT.SATPKS(25).NSAT.EXPNAM(5)
      COMMON /CSATLS/SATLST(25).SATINX
      COMMON /MODES/SGMODE(3).SHMODE(3).GMODE.HMODE
      COMMON /ERRORS/RMASS.ERRMAX.NERR.SIGERR.SECOND(411)
      COMMON /DUBFLG/IDUB
      COMMON /AJACNT/AJACEN(2)
      COMMON /DUBBUG/IDEBUG
      COMMON /CHEMPO/QDUBLT.QNORM.QSATS.NLINE
      COMMON. /QUIXOT/IQUIT
      COMMON /LSQ/PASS1
      COMMON /CRATMX/RATMXX
      COMMON /IMAD/MAD
      COMMON /UNITS/ IRDR.IPTR1.IPTR2.IPTR3.IIN.IOUT
      EQUIVALENCE    (YVAL(1).YM5).(YVAL(2).YM4).(YVAL(3).YM3)
      EQUIVALENCE    (YVAL(4).YM2).(YVAL(5).YM1).(YVAL(6).Y0)
      EQUIVALENCE    (YVAL(7).YP1).(YVAL(8).YP2).(YVAL(9).YP3)
      EQUIVALENCE    (YVAL(10).YP4).(YVAL(11).YP5)
      DATA HIST/-1.0.-.6667.-.3333.0...3333..6667.1..1.3333.1.6667.2./
      DATA PERCNT/.0125/
      DATA GTHT/'GETHGT'/
      DATA BEDROK/40/
      I1=NS-13-14*((NS-14)/14)
      I2=NS-12-14*((NS-13)/14)
      I3=NS-11-14*((NS-12)/14)
      I4=NS-10-14*((NS-11)/14)
      I5=NS-9-14*((NS-10)/14)
      I6=NS-8-14*((NS-9)/14)
      I7=NS-7-14*((NS-8)/14)
      I8=NS-6-14*((NS-7)/14)
      I9=NS-5-14*((NS-6)/14)
      I10=NS-4-14*((NS-5)/14)
      I11=NS-3-14*((NS-4)/14)
      I12=NS-2-14*((NS-3)/14)
      I13=NS-1-14*((NS-2)/14)
      I14=NS-14*((NS-1)/14)
C      *** ONLY GO HERE ON FIRST PASS TO COLLECT ANY PEAKS
C      *** THAT MIGHT HAVE BEEN INVOLVED IN A DOUBLET WITH
C      *** THE LAST COMPONENT THAT WAS PROCESSED ALSO
C      *** ZERO OUT SATURATION PARAMETERS.
C STATEMENT BELOW HAS BEEN COMMENTED OUT BECAUSE OF CALL TO
C FUNCTION SUBROUTINE SCAN.......
C                             RTI - WFH 11/20/78
C     DIVIS = HIMASS - LOMASS
      IF ( NDIS .EQ. 2    .AND.    AJACEN(1) .EQ. 1 ) PASS1 = 0
C           MAKE SURE WE SET UP RIGHT IN BACLSQ IF WE DON'T SCAN FOR M10
      IF ( NDIS .NE. 1 ) GO TO 6
C      NDIS = 1 ON FIRST PASS
      NSAT=0
      SATMAX=0
      RMASS = 32000
      ERRMAX = 0
      NERR = 0
      SIGERR = 0
      QNORM = 1.0
      QDUBLT = 0
      QSATS = 0
C      HELLO JOEF DROMEY.....WHERE EVER YOU ARE...TAO
      DO 1 I = 1.411
1     SECOND(I) = 32767
      IF(OVLAP.EQ.0) GO TO 6
C      HERE IF OVERLAPPING ELUTANT AHEAD.
      GMX=0.0
      HMX=0.0
      QDUBLT = 1
C      IDUB IS FLAG FOR SETGHY
      IDUB = 1
      DO 5 J=1.14
      IF(GMX.LT.GNU(J)) GMX=GNU(J)
      IF(HMX.LT.HNU(J)) HMX=HNU(J)
5     CONTINUE
```

```
C        NORMALIZE MODELS TO 1.0
       GMX=1.0/GMX
       HMX=1.0/HMX
       DO 7 J=1.14
       GNU(J)=GNU(J)*GMX
       HNU(J)=HNU(J)*HMX
7      CONTINUE
6      CONTINUE
C        *** NOW START LOOPING THROUGH ALL THE SAVED MASSES
       NMS=NM(MN)
       DO 500 MQ=1.411
        IF ( NUSPEC(MQ) .NE. -1 ) GO TO 500
C        NUSPEC(M) NE -1 IF INTENSITIES FROM EARLIER ELUTANT HAVE
C        BEEN ALREADY COMPUTED.
       M=IND(MN.MQ)
        IMASS = M
C        M < 0 MEANS MASS ALREADY CONSIDERED AS A COMPONENT OF PREVIOUS
C        SPECTRUM.
        IF (M .LT. 0) GO TO 500
C         IF M = 0 THEN MASS WAS NOT DETECTED AS PEAKING IN TLOC. SO TRY
C        FOR POSSIBLE DOUBLET.
        IF(M.EQ.0) M = MQ
        MI=M+39
        Y0=SPEC(I8.M)
C THIS MOD MADE AT RTI BECAUSE NO POSSIBLE REASON WAS SEEN
C FOR HAVING AN EQ INSTEAD OF A GE IN
CTHIS SITUATION.
C                                WFH - RTI 6/14/78
        IF(Y0.GE.SATVAL) GO TO 120
C      IF A DUBLET IS DETECTED RUN ALL MASSES THROUGH DOUBLET RESOLVER
        IF(OVLAP.EQ.1) GO TO 130
C        IF IMASS IS ZERO .THEN
C        NO SINGLET CONTRIBUTION DETECTED IN TLOC.
        IF(IMASS . EQ. 0) GO TO 500
        IF ( NDIS .EQ. 1   .AND.   AJACEN(1) .EQ. 1 ) GO TO 500
C        DON'T PROCESS ANY PASS ONE SINGLETS(MN=M1) IF A SPECTRUM
C        IMMEDIATELY PRECEDING THE CURRENTLY DETECTED SPECTRUM
C        WAS DETECTED.SINCE THIS SPECTRUM IS FOR MN=M2.
C        AND WE DON'T WANT TO INCLUDE SINGLETS FROM PREVIOUS SPECTRUM
C        IN IT'S IMMEDIATE SUCCESSOR.
C             BUT. WE MUST PROCESS SATURATED PEAKS  IN ALL CASES.
        YM1=SPEC(I7.M)
        YP1=SPEC(I9.M)
        YM5=SPEC(I3.M)
        YP5=SPEC(I13.M)
        YM4 = SPEC(I4.M)
        YP4=SPEC(I12.M)
        YM2=SPEC(I6.M)
        YP2=SPEC(I10.M)
        YM3=SPEC(I5.M)
        YP3=SPEC(I11.M)
C        PUT MASSES ON MODE OF PEAK USING CUBIC INTERPOLATION
        CALL NEWTOP(YVAL.RATEST.XTOP.YTOP.4)
C        MAKE SURE PEAKS ARE RISING FASTER THAN BACKGROUND.
C        THIS IS A VERY MINIMAL APPROXIMATIO OF BACKGROUND.
        IF( (YTOP-YM5) .LT. 90  .AND.  (YTOP-YP5) .LT. 90 ) GOTO 500
C        *** APPLY MINIMUM RATE CRITERION FOR HIGH MASS
C        *** PEAKS THAT CAN'T BE ISOTOPE PEAKS. TH:S ELIMINATES
```

```
C       *** BACKGROUND PEAKS AT THE LEADING EDGE OF COMPONENTS
C       *** NOTE THE COMPUTED RATES FOR BACKGROUND PEAKS ARE
C       *** ARE ALWAYS LOW BECAUSE THEIR INTENSITY VARIATION
C       *** IS CAUSED BY A DIFFERENT PROCESS TO NORMAL MASS
C       *** CHROMATOGRAM PROFILES FOR REAL COMPONENTS.
      IF(RATEST.LT..095) GO TO 500
      IF(M .LT. 140) GO TO 25
      IF(RATEST .LT. RATMIN)  GO TO 500
C       *** CORRECT FOR FINITE SWEEP OFFSET
C       *** REJECT ODDBALL PEAK-MODES
25    CONTINUE
30    CONTINUE
C       *** HERE FOR EXCLUSION TEST ON PEAK TOP RELATIVE TO
C       *** HISTOGRAM MODE - IF 'IH' IS OUTSIDE LIMITS THEN
C       *** PEAK IS REJECTED.
      IF(XTOP .LE. -1.5  .OR.  XTOP .GE.  1.5) GO TO 500
C      THE ABOVE REJECTS ANY PEAKS WITH FUNNY SHAPES, OR MODES NOT
C      IN HISTOGRAMMIC WINDOW.
      MI=M+39
      MI=M+39
C FUNCTION SUBROUTINE CALL HAS BEEN SUBSTITUTED FOR ARITHMETIC
C IN ORDER TO ACCOMODATE MAGNETIC INSTRUMENTS...
C                                    RTI - WFH 11/20/78
C     TSHIFT = (MI - LOMASS)/DIVIS
      TSHIFT = SCAN(MI)
     .XTOP = XTOP + TSHIFT
      DO 600 IJ = 1,10
      IF(XTOP .LT. HIST(IJ)) GO TO 700
600   CONTINUE
700   IH = IJ + (NDIS-1) * 3  -  4
C      IH IS IN ONE THE THREE SETS +-3 TO 60. +0 TO 90. +3 TO 120
C           AS A FUNTION OF NDIS = 1.2.3
C      SEE IF HISTOGRAM INDEX IS OUT OF RANGE OF DETECTED SPECTRUM
C      **** BELOW USED TO SAY IH.GE.NLWCUT*****12/19 CHANGED
      IF(IH.LT.NLWCUT.OR.IH.GT.NUPCUT) GO TO 500
C      WE NOW SEE IF THE DATA WILL PASS OVER A FLOOR
C      BELOW YTOP.
      FLUSH = PERCNT * YTOP
      BUBBLE = PERCNT * G(5)
C      GET  MAX+ FLUSH,BUBBLE @
      IF ( FLUSH .LT. BUBBLE ) FLUSH = BUBBLE
C      GET MIN+ 70,FLUSH @
      IF ( 70 .LT. FLUSH ) FLUSH = 70
C      GET MAX + BEDROK,FLUSH @
      IF ( BEDROK .GT. FLUSH ) FLUSH = BEDROK
      FLOOR = YTOP - FLUSH
C                         *
C
C
C
C
C          YM3       YTOP              YP3
C     ................................................ + FLOOR
C      IF THE CURVE STAYS ABOVE THE FLOOR IT GETS FLUSHED
      DO 110 II = 3,9
      IF ( YVAL(II) .LE. FLOOR )    GO TO 111
110   CONTINUE
      GO TO 500
```

```
C       ....     AND AWAY WE GOOOOOOOOOOOOOOOO ..... BLUB ...
111   CONTINUE
      CALL BLEED(YVAL(3).OITAR)
C
C THIS IS A TEMPORARY MOD IN ORDER TO GATHER STATICS ON
C THE INPUT PARAMETER ERRATIO - CODE 2
C                                         WFH RTI 9/31/78
C     ISPC=SPCNO
C     ITYP=2
C     WRITE(20.6250) ITYP.ISPC.OITAR.OITAR
C6250 FORMAT(A4.A4.16X.A4.4X.A4)
      IF ( OITAR .LT. RATMXX ) GO TO 500
C      THE ABOVE FITS A LINE TO YM3 THRU YP3. AND
C      RETURNS OITAR = (MEAN SQUARE ERR)/(AVERAGE VALUE Y )
C      *** THROW AWAY MASS INDEX SO THAT IT CAN'T BE INCLUDED
C      *** IN SPECTRUM OF NEXT-DETECTED COMPONENT
      IND(MN.MO)= -1
C      PUT MASSES ON SAME TIME CO-ORDINATE AS MODEL
      CALL TIMESF(TSHIFT.YVAL.11.GMODE)
C      *** NOW COMPUTE AND REMOVE LINEAR BACKGROUND BY
C      *** LEAST SQUARES METHOD USING MODEL PEAK - THE
C      *** CORRECTED INTENSITY IS RETURNED AS "PKHGT"
      CALL BACLSQ(PKHGT)
C       PEAKS OF INTENSITY < 10 ARE CONSIDERED AS ARTIFACTS OF
C      THE SINGLET RESOLVER
      IF( PKHGT .LT. 10 ) GO TO 500
      NUSPEC(M)=PKHGT
      GO TO 500
C      *** HERE FOR HANDLING SATURATED PEAKS BY RATIO
C      *** APPROACH THAT USES NON - SATURATED VALUES OF PEAK
C      *** TOGETHER WITH THE MODEL PEAK FOR THE COMPONENT TO
C      *** COMPUTE THE AMPLITUDE (LESS BACKGROUND) OF THE
C      *** SATURATED PEAK. JUMP OVER IF SATURATED PEAK HAS BEEN
C      *** PROCESSED IN AN EARLIER PASS BY "GETHGT"
120   CONTINUE
      NUSE  = NUSPEC(M)
      IF( NUSE .LT. 0  .AND.  NUSE .NE. -1 ) GO TO 500
C      NUSE < 0 IF THIS PEAK HAS BEEN PROCESSED IN EARLIER PASS
C      NOTE: IF NUSE = -1 IT HASN'T BEEN PROCESSED YET.
      IF(NSAT.GE.25) GO TO 500
      NSS=NS+2-NDIS
      CALL SATLSQ(NSS.M.RPKHGT.G.SPEC.SATVAL.MODE1.OVLAP.MODE2.PKHIST)
C      *** SAVE COMPUTED HEIGHT FOR LATER SO THAT IT CAN BE
C      *** NORMALIZED TO INTEGER SIZE
      IF ( IQUIT .EQ. 1 ) GO TO 500
C      IQUIT = 1 IF SATURATED VALUES MODE IS NOT "NEAR"
C      SPECTRUM.
C      IQUIT = 2 IF WE WANT TO SKIP PEAK BECAUSE IT DOESN'T
C      BELONG IN THIS SPECTRUM.IT WAS IN OVERLAPPING PREDECESSOR.
      IF ( IQUIT .NE. 2 ) GO TO 125
      NUSPEC(M) = -2
      GO TO 500
125   CONTINUE
      NSAT=NSAT+1
      SATPKS(NSAT)=RPKHGT
C      RPKHGT IS < 0 IF SUPERSATURATION ENCOUNTERED
      NUSPEC(M)= -NSAT - 2
C      SAVE MAXIMUM SATURATED VALUE ENCOUNTERED FOR NORMALIZATION
```

```
      IF(SATMAX.GT.RPKHGT) GO TO 500
      SATMAX=RPKHGT
      GO TO 500
C      *** HERE FOR HANDLING DOUBLETS IN THE SET
130   CONTINUE
C      RUN ALL MASSES THROUGH DOUBLET RESOLVER IF OVLAP=1.IE.
C      DOUBLET SITUATION DETECTED.
      IGPOS=MODE1
      IHPOS=MODE2
C      NSD=NS+3-NDIS
      IF(NDIS .NE. 2) GO TO 500
C      ONLY GO THROUGH DOUBLET RESOLVER WHEN WE ARE LOOKING AT
C      "SPECTRUM 7".
      CALL SETGHY(NS,M,GNU,HNU,IGPOS,IHPOS,IDUB,GMX)
C      DO NOT MARK MASS AS ALREADY FOUND "CONTRIBUTION" BY SETTING
C           IND(MN,MD) = 0.
C      SINCE THE DUBLET RESOLVER SAVES THE LARGEST CONTRIBUTION IN
C      THE ARRAY NUSPEC(M).
500   CONTINUE
C      *** HERE ONLY ON THIRD PASS AND ONLY THEN IF ANY
C      *** SATURATED PEAKS HAVE BEEN FOUND. NEED TO NORMALIZE
C      *** SO THAT EXTRAPOLATED VALUES CAN BE SAVED AS INTEGERS
      IF( NSAT .EQ. 0   .OR.   NDIS .NE. 3 ) RETURN
C      SET LOWER BOUND ON SATINX.THIS ARRAY IS USED IN SATLSQ.
      J = 26
      DO 723 I = 1,25
      J = J - 1
      IF ( SATLST(J) .NE. 0 ) GO TO 750
723   CONTINUE
750   CONTINUE
      SATINX = J
      AMXNU=0.0
C      GET MAXIMUM VALUE IN SPECTRUM
      DO 503 J=1,411
      IF(NUSPEC(J).GT.AMXNU) AMXNU=NUSPEC(J)
503   CONTINUE
C      SATMAX WILL BE 0 IF ONLY SUPERSATURATION ENCOUNTERED..
      IF ( SATMAX .EQ. 0 ) SATMAX = SATVAL
      IF(AMXNU.GT.SATMAX) SATMAX=AMXNU
      AMXNU = SATMAX
C      NORMALIZE TO 32000
      SATMAX=32000.0/SATMAX
C      NUSPEC(J) < 0 IF SATURATED VALUE FOR THAT MASS
      DO 530 J=1,411
      IF(NUSPEC(J)) 510,530,520
C      HERE IF SATURATED VALUE
510   IF ( NUSPEC(J) .EQ. -1 ) GO TO  530
      IF ( NUSPEC(J) .EQ. -2 ) GO TO  530
      NC = -NUSPEC(J) - 2
C      SATPKS(NC) < 0   IF SUPER SATURATION.
      IF ( SATPKS(NC) .LT. 0 ) SATPKS(NC) = AMXNU
      NUSPEC(J)=SATPKS(NC)*SATMAX
      QSATS = QSATS + 1
C      *** USE C13 ISOTOPE TO ASSIGN SATURATED PEAK TO THE
C      *** APPROPRIATE COMPONENT WHEN THERE ARE TWO COMPONENTS
C      *** CLOSE TOGETHER
C      OVLAP = 0 MEANS NO ELUTANT DETECTED AHEAD AND CLOSE
      IF( OVLAP .EQ. 0 )  GO TO 530
```

```
C       NUSPEC(J+1) NOT EQUAL ZERO MEAN ISOTOPE PRESENT IN THIS
C       ELUTANT
      IF ( NUSPEC(J+1) .EQ. 0 ) GO TO 511
      IF ( NUSPEC(J+1) .EQ. -1 ) GO TO 511
      GO TO 530
511   NUSPEC(J) = 0
      MY = J + 39-MAD
      QSATS = QSATS - 1
      IF ( IDEBUG .EQ. 0 ) GO TO 530
      WRITE(IPTR2.1000) MY
      GO TO 530
520   NUSPEC(J)=NUSPEC(J)*SATMAX
530   CONTINUE
      IF ( QSATS .GE. 1 ) QNORM = SATMAX
      RETURN
C3222  FORMAT(//' AJACENT FLAG ='.I3//)
1000  FORMAT(/.' MASS '.I3.' HAS NO C13 ISOTOPE PEAK './/)
      END
      SUBROUTINE KAPUT( SPECNO.TOTION.TICRAT.ABMAX )
      REAL TOTION.ABMAX
      INTEGER SPECNO.TICRAT
      COMMON /SLOTNO/NSLOT
      COMMON /CFSTPG/FSTPGE
      COMMON /CHEMPO/QDUBLT.QNORM.QSATS.NLINE
      COMMON /UNITS/ IRDR.IPTR1.IPTR2.IPTR3.IIN.IOUT
      INTEGER  QDUBLT.QSATS.NLINE.ITOTS.NSLOT.FSTPGE
      REAL     QNORM
      REAL   YES1/' './.YES/'YES './.YNOY/' NO '/
C      THIS ROUTINE DOES THE CHEMIST'S PRINTOUT
C      QDUBLT              1 IF THE SPECTRA IS FIRST ELUTANT OF DUBLET.
C      QSATS             NO. OF SATURATED PEAKS IN SPECTRUM.
C      NLINE             LINE COUNT FOR EACH PAGE.
C      SPECNO              SPECTRUM NUMBER
C      TOTION              TOTAL ION CURRENT OF SPECTRUM.
C      TICRAT              CLEANUP RATIO(SEE OUTDAT FOR DETAILS).
C      ABMAX             MAXIMUM AMPLITUDE OF SPECTRUM.
      NSLOT = NSLOT + 1
      IF(NSLOT.EQ.1)WRITE(IPTR3.1000)
      YES1=YES
      IF ( QDUBLT .EQ. 0 )YES1=YNOY
      WRITE(IPTR3.2000) NSLOT.SPECNO.TOTION.TICRAT.ABMAX.YES1.QSATS
      RETURN
1000  FORMAT('  SLOT  SPEC.*       TOTION  TICRAT  MAX. AMP   DOUBLET?
     *SAT. PEAKS '//)
2000  FORMAT(3X.I3.5X.I5.3X.F8.0.3X.I3.4X.F8.0.4X.A4.6X.I3)
      END
      SUBROUTINE LKAHED(NS.M2.M3.OVLAP.G.H)
C
C      THIS COMPUTER PROGRAM WAS DEVELOPED WITH THE FUNDING SUPPORT
C      FROM THE NATIONAL INSTITUTES OF HEALTH (GRANTS RR-612 AND
C      GM-20832) AND THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
C      (GRANT NGR-05-020-632).
C
C
C      *** THIS ROUTINE USES THE SAME TECHNIQUES AS 'TLOC'
C      *** TO LOOK BEYOND THE PROCESSING WINDOW AND DETECT
C      *** IF THERE ARE ANY NEW COMPONENTS NEARBY. THIS IS
C      *** IMPORTANT FOR COMPENSATING FOR OVERLAP
```

```
C       *** INTERRACTIONS USING THE DOUBLET RESOLVER.
C       *** A MODEL PEAK IS SELECTED FOR ANY COMPONENT THAT
C       *** IS COMING UP DIRECTLY. THE RELATIVE DISTANCE
C       *** BETWEEN NEIGHBORING COMPONENTS IS ESTABLISHED
C       *** ACCURATELY FROM THE LOCAL HISTOGRAM PROFILES.
C
      REAL    ISVP8. ISVP9
      INTEGER NAME.RDATE.BITMAP.GCTIME.LIBMAP.INTEG.LOMASS
      INTEGER HIMASS.POINTS.TAPCNT.MAXSP.LSTCNT.DEVICE.SPCNO
      INTEGER SPEC, IND. IR.BK.NM.NTM.BUBBLE
      INTEGER GPEAK.SPARE1.SPARE2.SPARE3.SPARE4.SPARE7
      REAL    SPARE5.SPARE6
      INTEGER NUSPEC.NMASS.OVLAP
      INTEGER NSP2.NSP3.NT2.NT3.SCNDPK.NERR.SECOND
      REAL NXTRTE.RASS.ERRMAX.SIGERR
C
      DIMENSION SCNDPK(3.11).NXTRTE(3).H(11).G(9)
      COMMON PKHIST(20).NPHIST(20)
      COMMON YM5.YM4.YM3.YM2.YM1.Y0.YP1.YP2.YP3.YP4.YP5
      COMMON GM4(9)
      COMMON SG.SG2.SIG.SI.SI2.AN.AA.BB.CC.DD
      COMMON NAME(32).RDATE(5).BITMAP(64).GCTIME
      COMMON LIBMAP(17), INTEG.LOMASS.HIMASS.POINTS, TAPCNT
      COMMON MAXSP.LSTCNT.DEVICE(5).SPCNO
      COMMON SPEC(14.411), IND(3.411).NUSPEC(411).NMASS
      COMMON RMAX(3), IR(3).BK(3).NM(3).NTM(3).GPEAK(3.11)
      COMMON SPARE1.SPARE2.SPARE3.SPARE4.SPARE5.SPARE6.SPARE7
      COMMON RATMIN.RATEMX.SIGY0.SIGNXT.SIGLST
      COMMON ISVN8. ISVN9. ISVP8. ISVP9. SIGMA7. SATVAL. SATMAX
      COMMON MODE1.MODE2.NLWCUT.NUPCUT.SATPKS(25).NSAT.EXPNAM(5)
      COMMON /CAHEA/AHEAD(5)
      COMMON /ERRORS/RASS.ERRMAX.NERR.SIGERR.SECOND(411)
      COMMON /MODES/ SGMODE(3).SHMODE(3).GMODE.HMODE
      COMMON /IFF/JFLAG(3)
      COMMON /PARAMS/MINTIC.MINNY.NTHIRD.RATM
      COMMON /MASSES/ISGMAS(3), ISHMAS(3). IGMASS. IHMASS
      COMMON /UNITS/ IRDR. IPTR1. IPTR2. IPTR3. IIN. IOUT
      INTEGER ONE.TWO.THREE.FOUR.FIVE.SIX.UNION.AHEAD
      DIMENSION IFLAG(3)
C
C
      REAL*8 ILKA
      DATA ONE/ 1/.TWO/2/.THREE/3/.FOUR/4/.FIVE/5/.SIX/6/
      DATA ILKA/'LKAHED'/
C      ************************************************************
      DO 3 I=12.20
      NPHIST(I)=0
      PKHIST(I)=0
3     CONTINUE
      DO 4 I=1.25
4     SATPKS(I)=0.0
      NSHIFT=0
      NSP2=NS+1
C      REINITIALIZE ARRAY FOR SAVING SPECTRUM
      DO 5 M=1.411
5     NUSPEC(M)= -1
      UNION = 0
      IF(OVLAP.EQ.0) GO TO 7
```

```
C        OVLAP SET MEANS PRECEDING SPECTRUM CONTRIBUTED TO THIS ONE.
        OVLAP=0
        UNION = 1
C        UNION IS SET TO INDICATE WE WILL UNION SPECTRUM N-1'S
C        CONTRIBUTIONS WITH SPECTRUM N ONLY IF SPECTRUM N
C        DOESN'T HAVE AN OVERLAPPING ELUTANT.
7       CONTINUE
        CALL NXTMAX(SPEC.NSP2.RNS2.YNS2.NT2.SATVAL
       &    .SCNDPK.NXTRTE.NSHIFT.IFLAG)
        NSHIFT=1
        NSP3=NS+2
        CALL NXTMAX(SPEC.NSP3.RNS3.YNS3.NT3.SATVAL
       &    .SCNDPK.NXTRTE.NSHIFT.IFLAG)
        NSHIFT=2
        NSP4=NS+3
        CALL NXTMAX(SPEC.NSP4.RNS4.YNS3.NT4.SATVAL
       &    .SCNDPK.NXTRTE.NSHIFT.IFLAG)
C        SCAN FROM   MODE1 + NTHIRD   TO 13. SINCE THE CURRENT SPECTRUM
C        EFFECTS HIST(2-8)  WITH MODE1 VARYING FROM 4 TO 6.
C        LOOKING FOR SPECTRUM IN +MODE1+NTHIRD.9 0
        NL=MODE1 + NTHIRD
C        SEE IF WE HAVE TO SCAN TOO CLOSE ( +7.8.90)
        IF ( NL .GE. 10 ) GO TO 10
C        SO NL .LE. 9
        IDXL = NL
        IDXU = 9
        NU = 10
        IAHEAD = 2
        CALL DETEKT(PKHIST.NPHIST.OVLAP.NLWCUT.NUPCUT.NL.NU.IDXL  .IDXU.MO
       *DE2)
        IF ( OVLAP .EQ. 1 ) GO TO 26
C        SCAN +9.130 FOR ELUTANT IN +10.120
        NL = 9
10      IF ( NL .GE. 13 ) GO TO 20
        IAHEAD = 3
        NU=13
        IDXL=10
        IDXU=12
        CALL DETEKT(PKHIST.NPHIST.OVLAP.NLWCUT.NUPCUT.NL.NU.IDXL
       &    .IDXU.MODE2)
        IF(OVLAP.EQ.1) GO TO 26
C        SEE IF SPECTRUM AHEAD IN +13.150
20      NL = 12
        NU =16
        IDXL = 13
        IDXU = 15
        IAHEAD = 4
        CALL DETEKT(PKHIST.NPHIST.OVLAP.NLWCUT.NUPCUT.NL.NU.IDXL.IDXU.MODE
       *2)
        IF ( OVLAP .EQ. 1 ) GO TO 26
C        SEE IF SPECTRUM AHEAD IN +16.180
        NL = 15
        NU = 19
        IDXL = 16
C       IDXL=18
C THIS SEEMED AN OBVIOUS TYPO. SO IDXL HAS BEEN CHANGED
C TO IDXU AS IN THE STMT BELOW
C                                WFH RTI 9/31/78
```

```
      IDXU = 10
      IAHEAD = 5
      CALL DETEKT(PKHIST.NPHIST.OVLAP.NLWCUT.NUPCUT.NL.NU.IDXL.IDXU.MODE
     *2)
      IF ( OVLAP .EQ. 0 ) GO TO 30
26    CONTINUE
      AHEAD(IAHEAD) = 1
C      SET SPECTRUM AHEAD FLAG
30    CONTINUE
      RC0 = .5
      RATEMX = RMAX(M2)
      RATMIN = RATEMX * 0.25  * RC0
      GMODE=SGMODE(M2)
      JFLAG1 = JFLAG(M2)
      IGMASS = ISGMAS(M2)
C      STORE MODEL PEAK FOR THIS SPECTRUM
      DO 2000 IJ=1.9
2000  GM4(IJ) = GPEAK(M2.IJ+1)
      IF(OVLAP.EQ.0) GO TO 95
C      HERE FOR OVERLAPPING SPECTRUM DOWN THE PIPE
      MODEH=6+(MODE2-1)/3
C      MODEH IS RELATIVE SPECTRUM NUMBER OF OVER LAPPER.
C LABEL 40 HAS BEEN PUT TO BETTER USE....
C                             RTI WFH 11/20/78
      IF(MODEH-10) 50.60.70
50    INDXH=1
      GO TO 80
60    INDXH=2
      GO TO 80
70    INDXH=3
80    CONTINUE
      SH=0.0
C      STORE MODEL H
C THIS MOD INSERTED TO HELP WITH POOR INITIALIZATION
C OF H-PEAK PARAMETERS.......
C                             RTI WFH 11/20/78
      IF(ISHMAS(INDXH).NE.-1)GO TO 40
      OVLAP = 0
      RETURN
   40 IF ( MODEH .NE. 8 ) GO TO 89
C      IF SPECTRUM ELUTING AT SLOT 8.GIVE IT DETECTED THERE.
      DO 88 I = 1.11
      H(I) = GPEAK(M3.I)
88    SH = SH + H(I)
      IHMASS = ISGMAS(M3)
      HMODE = SGMODE(M3)
      IFLAG(INDXH) = JFLAG(M3)
      GO TO 91
89    CONTINUE
      HMODE = SHMODE(INDXH)
      IHMASS = ISHMAS(INDXH)
      DO 90 I=1.11
      H(I) = SCNDPK(INDXH.I)
      SH=SH + H(I)
90    CONTINUE
91    CONTINUE
C      SEE IF WEIRD H MODEL PEAK.
      IF ( SH .LE. 0 ) OVLAP = 0
      IF ( SH .LE. 0 ) GO TO 95
      CALL TIMESF( HMODE.H.11.GMODE )
C      ALIGN H-MODEL TO G-MODEL'S TIME CO-ORDINATE.
C          .. ANOTHER VOID .. EH       DOGEN ..
95    CONTINUE
C      SEE IF PREVIOUS SPECTRUM OVERLAPPED WITH THIS ONE.
      IF ( UNION .EQ. 0 ) GO TO 27
      BUBBLE = .01 * GM4(5)
C      UNION = 0 IF PREVIOUS ELUTANT NOT OVERLAPPING WITH THIS ONE.
C      WE WANT TO MAKE SURE THOSE "SMALL CONTRIBUTIONS"
C      WHEN CALCULATING THE CURRENT  SPECTRUM USING
C      THE DUBLET RESOLVER IN CONJUNTION WITH THE PREVIOUS
C       OVERLAPPING ELUTANT GET USED IF WE ARE AGAIN IN
C      AN OVERLAPPING SITUATION.
      IF ( OVLAP .EQ. 0 ) GO TO 29
C      ONLY DO BELOW IF DOUBLET IS AGAIN A DOUBLET
      DO 28 M = 1.NMASS
28    IF ( SECOND(M) .LE. BUBBLE ) NUSPEC(M) = 0
      GO TO 27
C      OVLAP = 1. SO AGAIN IN OVERLAPPING SITUATION. DON'T UNION
C      N-1 WITH N. RATHER. CALCULATE N'S SPECTRUM(EXCEPT FOR
C      ZERO CONTRIBUTIONS FROM N-1) USING DUBLET RESOLVER.
29    CONTINUE
C      OVERLAP = 0 SO NO DUBLET SITUATION.THEREFORE UNION.
C      SINCE UNION = 1 WE'LL  UNION SPECTRUM N-1'S CONTRIBUTIONS
C      WITH OUR CURRENT SPECTRUM. N.
      DO 6 M = 1.411
      IF ( SECOND(M) .LE. 10 )    SECOND(M) = 0
6     IF ( SECOND(M) .NE. 32767 ) NUSPEC(M) = SECOND(M)
27    CONTINUE
      DO 2400 IJ=1.9
2400  G(IJ)=GM4(IJ)
      CALL LKAPRT(7.NPHIST.PKHIST)
C      THE ABOVE PRINTS THE SECOND HALF OF THE HISTOGRAMS.
      IF(JFLAG1 .EQ. 1) CALL LKAPRT(FIVE.I.R)
C      THIS HAPPENS IF THE MODEL IS FLAGGED AS FUNNY IN TLOC
      CALL LKAPRT(THREE.IGMASS.G)
C      PRINT MODELS
      IF(OVLAP .EQ. 0) RETURN
      CALL LKAPRT(ONE.I.R)
C      THE ABOVE  PRINTS 'OVERLAPPING SPECTRUM FOUND' MSG.
      IF(MODEH.GE.9.AND.MODEH.LE.11) GOTO 401
C      THE BELOW WRITES MODEH AS A WARNING THAT THE SECOND
C          ELUTANT IS VERY CLOSE
      CALL LKAPRT(TWO.MODEH.H)
401   CONTINUE
C      MODEH = 12 ONLY IF NEXT MAX IS IN SLOTS 19.20 OF HISTOS.
      CALL LKAPRT( FOUR.IHMASS.H(2) )
C      SEE IF FLAKEY MODEL WAS USED FOR INTERPOLATION
      IF(IFLAG(INDXH) .EQ. 1) CALL LKAPRT(SIX.I.H)
      RETURN
      END
C      THE CALL STATEMENT ALLOWS FOR PASSING OF ARRAYS
C      REALS AND INTEGERS.
C      THIS COMPUTER PROGRAM WAS DEVELOPED WITH THE FUNDING SUPPORT
C      FROM THE NATIONAL INSTITUTES OF HEALTH (GRANTS RR-612 AND
C      GM-20832) AND THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
C      (GRANT NGR-05-020-632).
```

```
      SUBROUTINE LKAPRT(N.IAR.AR)
      INTEGER N.IONE.NINE.MAD.IIAR(20)
      INTEGER IAR(1)
      REAL AR(1)
      COMMON /DUBBUG/IFLAG
      COMMON /IMAD/MAD
      COMMON /UNITS/ IRDR.IPTR1.IPTR2.IPTR3.IIN.IOUT
      DATA IONE.NINE/1.9/
C      IFLAG IS RESET IF NO PRINTING IS TO BE DONE/
      IF (IFLAG.EQ.0) RETURN
      IIAR(1)=IAR(1)-MAD
      IF(IIAR(1).LT.0)IIAR(1)=0
      GO TO (80.90.100.110.120.130.140).N
C       THE ABOVE IS A CASE STATEMENT FOR EXECUTING PRINT
80    WRITE(IPTR2.580)
      RETURN
C   THE ABOVE  WRITE IS CALLED FROM LKAHED(BELOW ALSO)
90    WRITE(IPTR2.590) IIAR(1)
      RETURN
100   WRITE(IPTR2.1000) (AR(I). I = IONE.NINE).IIAR(1)
      RETURN
110   WRITE(IPTR2.1100) (AR(I). I = IONE.NINE).IIAR(1)
      RETURN
120   WRITE(IPTR2.1200)
      RETURN
130   WRITE(IPTR2.1300)
      RETURN
140   WRITE(IPTR2.1500)
      DO 5 K=2.18
      IIAR(K)=IAR(K)-MAD
      IF(IIAR(K).LT.0)IIAR(K)=0
    5 CONTINUE
      WRITE(IPTR2.530) (IIAR(I). I = 10.18 )
      WRITE(IPTR2.535) ( AR(I).  I = 10.18 )
C      WRITES SECOND HALF OF HISTOGRAMS.
      RETURN
580   FORMAT(/.' OVERLAPPING SPECTRUM FOUND')
590   FORMAT(' OVERLAPPING ELUTANT IS CLOSE. RELATIVE SPEC. =',I4)
1000  FORMAT(/.' MODEL USED FOR SINGLET RESOLUTION './.9F9.0.
     -//.' MASS OF MODEL =',I3)
1100  FORMAT(/'  MODEL USED FOR DOUBLET RESOLUTION './.9F9.0.
     -//' MASS OF MODEL = '.I3./.'   NOTE: THIS MODEL HAS BEEN'
     -.' INTERPOLATED TO THE MASS OF THE SINGLET MODEL')
1200  FORMAT(' *** WARNING *** THIS SPECTRUM HAS A SUSPECT G-MODEL '
     -.'PEAK'/.'   POSSIBLE BACKGROUND PEAK USED AS MODEL'/
     -.'   EXAMINE MODEL USING  LOOK  IF YOU ARE WARY'/)
1300  FORMAT(' *** WARNING *** THIS SPECTRUM HAS A SUSPECT H-MODEL '
     -.'PEAK'/.'   POSSIBLE BACKGROUND PEAK USED AS MODEL'/
     -.'   EXAMINE MODEL USING  LOOK  IF YOU ARE WARY'/)
535   FORMAT(1X.9F9.0)
530   FORMAT(/.' PEAK HIST  './.8X.9I6//' ION CURRENT HIST ')
1500  FORMAT(//' HISTOGRAM POSITIONS 10 - 18 ...'/)
      END
      SUBROUTINE NEWTOP(Y.RATE.XTOP.YTOP.ISTART)
      REAL    RATE.XTOP
      DIMENSION    Y(11)
      INTEGER   ISTART
C
```

```
C       THIS COMPUTER PROGRAM WAS DEVELOPED WITH THE FUNDING SUPPORT
C       FROM THE NATIONAL INSTITUTES OF HEALTH (GRANTS RR-612 AND
C       GM-20832) AND THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
C       (GRANT NGR-05-020-632).
C
C       THIS SUBROUTINE INTERPOLATES A PEAK TO THE MODE OF A PARABOLA
C       THRU THE MIDDLE 5 POINTS. THE MODE IS FOUND USING A PARABOLIC
C       LEAST SQUARE FIT THROUGH THE MIDDLE FIVE Y'S . THEN
C       SOLVING FGR THE X DISPLACEMENT OF THE MODE  FROM 0.
C
C            SO. IF Y = A + B * X + C * X**2
C
C       X = -B/2C IS THE DISPLACEMENT.
C                      Y
C                      *
C                      :          *
C              *         :
C                      :
C                      :                 *
C         *              :
C       ...................:........................ X
C                      0     -B/2C
C
C       THE INTERPOLATION IS DONE BY A CUBIC FIT USING  LAGRANGE'S
C       METHOD.I.E.. GIVEN THE 4 POINTS  X0.X1.X2.X3. WE FIND L(X) AS
C       FOLLOWS:
C
C            L(X)=+(X-X1)(X-X2)(X-X3)/((X0-X1)(X0-X2)(X0-X3))@Y0 +
C                 +(X-X0)(X-X2)(X-X3)/((X1-X0)(X1-X2)(X1-X3))@Y1 +
C                 +(X-X0)(X-X1)(X-X3)/((X2-X0)(X2-X1)(X2-X3))@Y2 +
C                 +(X-X0)(X-X1)(X-X2)/((X3-X0)(X3-X1)(X3-X2))@Y3
C
C            WITH THE X'S INTEGRALLY SPACED.
C
C       IN EACH CASE WE INTERPOLATE FOR A NEW MIDDLE POINT.OR
C       L(X1+DX). SO. THE ABOVE EQUATION REDUCES TO
C
C            L(X1+DX) = C0 * Y0 + C1 * Y1 + C2 * Y2 + C3 * Y3
C
C          C0 = DX * (DX-1)(DX-2)/-6
C       C1 = (DX+1)(DX-1)(DX-2)/2
C       C2 = (DX+1)(DX)(DX-2)/-2
C       C3 = (DX+1)(DX)(DX-1)/6
C
C       WHERE DX = -B/2C
C
C THIS ROUTINE HAS BEEN CHANGED SO THAT Y(I) ARE UNCHANGED
C AND A SMOOTHED RATE IS CALCULATED FROM THE WORK ARRAY Z(I)
C                                     WFH RTI 9/31/78
      INTEGER I
      REAL    C1,C2.C3.K1,K2.K3
      REAL    YI.YI1
      REAL Z(11)
      DIMENSION YOLD(11)
C
C      SAVED OLDIES
C
      XTOP = 2
      XB = 2 * Y(4) - Y(5) - 2 * Y(6) - Y(7) + 2 * Y(8)
      IF(XB .EQ. 0) RETURN
      XTOP = .7 * (2 * Y(4) + Y(5) - Y(7) - 2 * Y(8))/XB
      DX = XTOP
      K1 = DX - 1
      K2 = DX - 2
      K3 = DX + 1
      C1= (DX * (K1) * (K2))/(-6)
      C2= ((K3) * (K1) * (K2))/2
      C3= ((K3) * (DX) * (K2))/(-2)
      C4= ((K3) * (DX) * (K1))/6
C
C      INTERPOLATE
C
      YTOP = Y(5) * C1 + Y(6) * C2 + Y(7) * C3 + Y(8) * C4
      IS = ISTART - 1
C   ...   SMOOTH PEAK ALITTLE BEFORE COMPUTING RATE.
      Z(1)=(Y(1)+.1*Y(2))/1.1
      DO 275 I=2.10
      Z(I)=(Y(I)+.1*(Y(I-1)+Y(I+1)))/1.2
275   CONTINUE
      Z(11)=(Y(11)+.1*Y(10))/1.1
C SEE MOD NOTE AT BEGINNING WFH RTI 9/31/78
      DO  250 I = IS.8
250   IF ( Z(I) .EQ. 0 ) Z(I) = 1
C      CALCULATE RATE TO BE AN APPROXIMATION OF
C
C              INTEGRAL(DY/Y)       ...HIGH RATE FOR SHARP PEAKS.
      RATE = 0
C SEE MOD NOTE AT BEGINNING WFH RTI 9/31/78
      DO 300 I = ISTART.6
300   RATE = RATE + (Z(I) - Z(I-1))/Z(I-1)
      DO 400 I = 6.7
400   RATE = RATE + (Z(I) - Z(I+1))/Z(I+1)
      RETURN
      END
      SUBROUTINE NXTMAX(SPEC.NS.RATEMX.YMAX.NTMAX.SATVAL
     &   .SCNDPK.NXTRTE.NSHIFT.IFLAG)
C
C      THIS COMPUTER PROGRAM WAS DEVELOPED WITH THE FUNDING SUPPORT
C      FROM THE NATIONAL INSTITUTES OF HEALTH (GRANTS RR-612 AND
C      GM-20832) AND THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
C      (GRANT NGR-05-020-632).1.ID2
C
      INTEGER SPEC.SCNDPK.IFLAG(3)
      INTEGER RDATE.BITMAP.GCTIME.HIMASS.POINTS.TAPCNT
      INTEGER NX.NTMAX.IYD.ID1.ID2
      DIMENSION SPEC(14.411)
      DIMENSION SCNDPK(3.11).NXTRTE(3)
      COMMON PKHIST(20).NPHIST(20)
      COMMON YM5(11)
      COMMON GM4.GM3.GM2.GM1.G0.GP1.GP2.GP3.GP4
      COMMON SG.SG2.SIG.SI.SI2.AN.AA.BB.CC.DD
      COMMON NAME(32).RDATE(5).BITMAP(64).GCTIME
      COMMON LIBMAP(17).INTEG.LOMASS.HIMASS.POINTS.TAPCNT
      COMMON /MODES/SGMODE(3).SHMODE(3).GMODE.HMODE
      COMMON /LFTRIT/ILEFT.IRIGHT
      COMMON /MASSES/ISGMAS(3).ISHMAS(3).IGASS.IHMASS
      EQUIVALENCE  (YM4,YM5(2))
      EQUIVALENCE (YM3,YM5(3)),(YM2,YM5(4)),(YM1,YM5(5))
      EQUIVALENCE (Y0,YM5(6)),(YP1,YM5(7)),(YP2,YM5(8))
      EQUIVALENCE (YP3,YM5(9)),(YP4,YM5(10)),(YP5,YM5(11))
C
      INTEGER LOCLBG,ILEFT
      REAL NXTRTE,SCAN
      REAL HIST(7),SLOPE,YM5MIN,YP5MIN
      DATA HIST/-.3333,0.,.3333,.6667,1.0,1.3333,1.6667/
C
C      *** ROUTINE TO LOOK AHEAD TO SEE IF THERE ARE ANY OTHER
C      *** SPECTRA IN THE LOCAL REGION OF CURRENT SPECTRUM
C
      I1=NS-13-14*((NS-14)/14)
      I2=NS-12-14*((NS-13)/14)
      I3=NS-11-14*((NS-12)/14)
      I4=NS-10-14*((NS-11)/14)
      I5=NS-9-14*((NS-10)/14)
      I6=NS-8-14*((NS-9)/14)
      I7=NS-7-14*((NS-8)/14)
      I8=NS-6-14*((NS-7)/14)
      I9=NS-5-14*((NS-6)/14)
      I10=NS-4-14*((NS-5)/14)
      I11=NS-3-14*((NS-4)/14)
      I12=NS-2-14*((NS-3)/14)
      I13=NS-1-14*((NS-2)/14)
      I14=NS-14*((NS-1)/14)
C
C TAKE OUT - THIS STATEMENT NO LONGER NECESSARY WITH
C SCAN SUBROUTINEINSERTED.
C                                        WFH RTI 9/31/78
C     DIVISO = HIMASS - LOMASS
      RATEMX=0.0
      NTMAX=0
      NPOS=NSHIFT+1
C      NOW LOOP THROUGH MASSES
C
      DO 60 M=1,411
C
      MI = M + 39
C
C      MI IS THE ACTUAL MASS NUMBER
C
      Y0=SPEC(I8,M)
C THIS MOD MADE AT RTI BECAUSE NO POSSIBLE REASON
C WAS SEEN FOR HAVING AN EQ INSTEAD OF A GE IN
C THIS SITUATION.
C                              WFH - RTI 6/14/78
      IF(Y0 .GE. SATVAL) GO TO 60
      YM1=SPEC(I7,M)
      IF(Y0 .LE. YM1) GO TO 60
C      PEAK TOP A Y0
      YP1=SPEC(I9,M)
      IF(YP1 .GE. Y0) GO TO 60
C      WE NOW HAVE Y0 AT TOP OF PEAK
40    YM2=SPEC(I6,M)
      YP2=SPEC(I10,M)
      YP3=SPEC(I11,M)
      YM3=SPEC(I5.M)
      YM4 = SPEC(I4.M)
      YP4=SPEC(I12,M)
      YM5(1)=SPEC(I3.M)
      YY5 = YM5(1)
C      WE HAVE ONLY 14 SPECTRA IN CORE AT A TIME SO:
C
C          NSHIFT  SPECTRA AVAILABLE
C                TO RIGHT OF MODE
C
C            0           5
C            1           4
C            2              3
      IF(NSHIFT-1) 41.42.43
41    YP5=SPEC(I13,M)
      GO TO 44
42    YP5 = 0.0
      GO TO 44
43    YP4 = 0.0
      YP5 = 0.0
44    CONTINUE
C
C      GET HERE TO QUALIFY AS A SINGLET FOR HISTOGRAMMING
      IF ( YM2 .GE. YM1 ) GO TO 60
      IF ( YM3 .GE. YM2 ) GO TO 60
      IF ( ILEFT .EQ. 3 ) GO TO 100
      IF ( YM4 .GE. YM3 ) GO TO 60
100   CONTINUE
      IF ( YP2 .GE. YP1 ) GO TO 60
      IF ( YP3 .GE. YP2 ) GO TO 60
      IF ( NSHIFT .EQ. 2 ) GO TO 200
      IF ( YP4 .GE. YP3 ) GO TO 60
200   CONTINUE
C      PUT Y0 AT TOP OF PEAK
      CALL NEWTOP(YM5.RATE.XTOP.YTOP,4)
C      CHECH FOR NARROW RISING PEAK (CHAR. OF BACKGROUND) AND.
C      CHECK FOR PEAK MODE OUT OF BOUNDS. IF SO THE DATA HAD SOME
C      INFLECTION POINTS.
      YYM5=YTOP - YY5
      IF ( NSHIFT .EQ. 0 ) YX5 = YP5
      IF ( NSHIFT .EQ. 1 ) YX5 = YP4
      IF ( NSHIFT .EQ. 2 ) YX5 = YP3
      YYP5 = YTOP - YX5
      IF ( YYP5 .LT. 30  .AND.  YYM5 .LT. 30 ) GO TO 60
      IF(XTOP .LE. -1.   .OR.  XTOP .GE. 1.) GO TO 60
C      CHECK FOR SATURATION
      IF ( YTOP .GE. SATVAL ) GO TO 60
      IF ( YM1 .GE. SATVAL ) GO TO 60
      IF ( YM2 .GE. SATVAL ) GO TO 60
      IF ( YM3 .GE. SATVAL ) GO TO 60
      IF ( YM4 .GE. SATVAL ) GO TO 60
      IF ( YP1 .GE. SATVAL ) GO TO 60
      IF ( YP2 .GE. SATVAL ) GO TO 60
      IF ( YP3 .GE. SATVAL ) GO TO 60
      IF ( YP4 .GE. SATVAL ) GO TO 60
C      SEE IF WE HAVE MAXIMUM RATE
      IF ( NSHIFT .GE. 1 ) GO TO 300
      SLOPE = (YP5 - YY5)/10.
      LOCLBG = YY5 + 5 * SLOPE
300   CONTINUE
C     TSHIFT = (MI - LOMASS)/DIVISO
C THIS SECTION CHANGED TO FUNCTION SUBROUTINE TO
C GAIN INSTRUMENT COMPATABILITY.
C                                         WFH RTI 9/31/78
      TSHIFT=SCAN(MI)
      IF(RATE.LT.RATEMX) GO TO 50
      IF((YTOP - YY5) .GE. 160) GO TO 5000
      IFLAG(NPOS) = 1
C      FLAG FUNNY MODEL POSSIBLE
      GO TO 5500
5000  IFLAG(NPOS) = 0
5500  CONTINUE
C      GET HERE TO QUALIFY AS POSSIBLE MODEL FOR SPECTRUM
      RATEMX=RATE
C
C      USE A LINEAR BACKGROUND FIRST CUT APPROXIMATION.
C
      MODBKG = SPEC(I1,M)
C      PICK MIN ↑ SPEC(IJ,M) : I1 .LE.  IJ  .LE. I5 0
      IF ( MODBKG .GT. SPEC(I2,M) ) MODBKG = SPEC(I2,M)
      IF ( MODBKG .GT. SPEC(I3,M) ) MODBKG = SPEC(I3,M)
      IF ( MODBKG .GT. SPEC(I4,M) ) MODBKG = SPEC(I4,M)
      IF ( MODBKG .GT. SPEC(I5,M) ) MODBKG = SPEC(I5,M)
      IF ( NSHIFT .GE. 1 ) LOCLBG = MODBKG
C      USE CONSTANT BACKGROUND IF NO RIGHT END POINT AVAILABLE
      DO 450 IJ = 1,11
      SCNDPK(NPOS,IJ) = YM5(IJ) - MODBKG
450   CONTINUE
      SHMODE(NPOS) = TSHIFT
C
C      SAVE TSHIFT FOR LATER USE
C
      ISHMAS(NPOS) =  MI
C      SAVE MASS OF MODEL PEAK
50    CONTINUE
      CONTRI = YTOP - LOCLBG
      IF ( CONTRI .LE. 0 ) GO TO 60
      TITTY = TSHIFT + XTOP
      DO 500  I=1,7
      IF (TITTY .LE. HIST(I))  GO TO 600
500   CONTINUE
600   CONTINUE
      IH = 7 + I + 3*NSHIFT
C
C      IH IS IN ONE OF THE THREE SETS  ↑8 TO 140,↑11 TO 170  OR
C            ,↑14 TO 200
C            AS A FUNCTION OF NSHIFT. NSHIFT = 0, 1 OR 2.
C
C      UPDATE HISTOGRAMS
C
      PKHIST(IH)=PKHIST(IH)+ CONTRI
      NPHIST(IH) = NPHIST(IH) + 1
C           ....JUST A MINOR VOID ..HUI NENG..
60    CONTINUE
C
C
```

```
      NXTRTE(HPOS) = RATEMX
C
      RETURN
      END
      SUBROUTINE OFFSET (G.H.GNU.HNU.GPOS.HPOS)
C
C      THIS COMPUTER PROGRAM WAS DEVELOPED WITH THE FUNDING SUPPORT
C      FROM THE NATIONAL INSTITUTES OF HEALTH (GRANTS RR-612 AND
C      GM-20832) AND THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
C      (GRANT NGR-05-020-632).
C
C
C      *** THIS ROUTINE TAKES THE 'GPOS' AND 'HPOS' INFO
C      *** WHICH CHARACTERIZES WHERE ON THE SPECTRUM SCALE
C      *** TWO OVERLAPPING COMPONENTS ARE LOCATED WITH
C      *** RESPECT TO ONE ANOTHER AND MAKES THE NECESSARY
C      *** DISPLACEMENTS OF THE TWO MODEL PEAKS (ONE FOR
C      *** EACH OF THE OVERLAPPING COMPONENTS) WITH
C      *** RESPECT TO EACH OTHER SO THAT THEY CAN
C      *** BE USED BY THE DOUBLET RESOLVING ROUTINE
C      *** IN CONJUNCTION WITH THE RAW DATA FOR
C      *** THE DOUBLETS THAT HAVE TO BE RESOLVED AND
C      *** BACKGROUND CORRECTED.
C
      DIMENSION G(9).H(11).GNU(18).HNU(18)
      INTEGER GPOS.HPOS.HNUT(18).IHSPEC.JLOWER
      REAL*8 OFST
      DATA OFST/'OFFSET'/
C
C      ***STRETCH OUT THE MODEL PEAKS A LITTLE
C
      DO 10 I=1.18
      GNU(I)=0
      HNU(I)=0
10    CONTINUE
C      PUT G AND H INTO LONGER ARRAYS SO THAT THEY CAN BE ALIGNED TO
C      SPECTRUM HISTOGRAMIC POSITIONS.
      DO 30 I=1.9
      GNU(I+2) = G(I)
30    CONTINUE
C      DON'T USE H(1). AND H(11).THESE ARE EXTRA FOR INTERPOLATION
C      IN LKAHED.
      DO 110 I=1.9
      HNUT(I) = H(I+1)
110   CONTINUE
C
C
C      ALIGN TO SPECTRUM OF H MODEL
C
      IHSPEC = 6 + (HPOS - 1)/3
C
C      6 .LE. IHSPEC .LE. 11
C
      JLOWER = IHSPEC - 4
      DO  1000  I= 1.9
      HNU(JLOWER) = HNUT(I)
      JLOWER = JLOWER + 1
1000  CONTINUE
```

```
C
C       NOW. HNU(IHSPEC) REPRESENTS THE MODE OF THE H-MODEL.  AND
C       GNU(7) IS THE MODE OF THE G-MODEL
C
      RETURN
      END
      SUBROUTINE OUTDAT(SRLIB.NNAME)
      INTEGER NNAME(18).MIN.MINI.MAX.MAXI
      INTEGER MAD
      INTEGER PDATE(5).PNAME(32).FLAG
      INTEGER POINTX.TAPCNX.MAXSX
      INTEGER OUTSLT.MAXAMP.EXPNAM.TICRAT
      INTEGER NAME.RDATE.BITMAP.GCTIME.LIBMAP.INTEG.LOMASS
      INTEGER HIMASS.POINTS.TAPCNT.MAXSP.LSTCNT.DEVICE.SPCNO
      INTEGER SPEC.IND.IR.BK.NM.NTM
      INTEGER GPEAK.SPARE1.SPARE2.SPARE3.SPARE4.SPARE7
      INTEGER NUSPEC.NMASS.OVLAP.SRLIB(5)
      INTEGER JM(411).JSP(411)
      INTEGER QDUBLT.QSATS.NLINE
      INTEGER NUZEN.IMOLWT
      REAL RMM(411).RMSP(411)
      REAL CLASS(4)/4*' '/.SUBCL(4)/4*' '/
      REAL      SPARE5.SPARE6
      REAL*4 TOTION.SIGMA7
      REAL      ISVP8.ISVP9
      REAL   QNORM.ABMAX
      REAL   SATSUM.FORM(5)
      REAL*8 OTDT/' OUTDAT '/.SOURCE/' '/
      DIMENSION G(9).H(9).GNU(18).HNU(18)
      COMMON PKHIST(20).NPHIST(20)
      COMMON YM5.YM4.YM3.YM2.YM1.Y0.YP1.YP2.YP3.YP4.YP5
      COMMON GM4.GM3.GM2.GM1.G0.GP1.GP2.GP3.GP4
      COMMON SG.SG2.SIG.SI.SI2.AN.AA.BB.CC.DD
      COMMON NAME(32).RDATE(5).BITMAP(64).GCTIME
      COMMON LIBMAP(17).INTEG.LOMASS.HIMASS.POINTS.TAPCNT
      COMMON MAXSP.LSTCNT.DEVICE(5).SPCNO
      COMMON SPEC(14.411).IND(3.411).NUSPEC(411).NMASS
      COMMON RMAX(3).IR(3).BK(3).NM(3).NTM(3).GPEAK(3.11)
      COMMON SPARE1.SPARE2.SPARE3.SPARE4.SPARE5.SPARE6.SPARE7
      COMMON RATMIN.RATEMX.SIGY0.SIGNXT.SIGLST
      COMMON ISVN8.ISVN9.ISVP8.ISVP9.SIGMA7.SATVAL.SATMAX
      COMMON MODE1.MODE2.NLWCUT.NUPCUT.SATPKS(25).NSAT.EXPNAM(5)
      COMMON /PARAMS/MINTIC.MINNY.NTHIRD.RATM
C     COMMON /SCR/ LIBBLK
      COMMON /ERRORS/ERRMIN.ERRMAX.NERR.SIGERR.IACERR(411)
      COMMON /PRNTFG/IPFLAG
      COMMON /MASSES/ISGMAS(3).ISHMAS(3).IGMASS.IHMASS
      COMMON /DUBBUG/IDEBUG
      COMMON /CHEMPO/QDUBLT.QNORM.QSATS.NLINE
      COMMON /IMAD/MAD
      COMMON /NA/PNAME
      COMMON /UNITS/ IRDR,IPTR1.IPTR2.IPTR3.IIN.IOUT
C THESE EQUIVALENCE STMTS. HAVE BEEN FOUND SUPERFLUOUS IN
C THE MULTIPLE RUN PROCESSING STRUCTURE OF CLEANUP
C AND HAVE BEEN DELETED . THE VARIABLES INSIDE OUTDAT
C WERE MODIFIED IN THE FOLLOWING MANNER:
C       EXPID.PDATE.AMP.LOMASX.HIMASX.LSTCNX.GCTIMX.
C       INTEGX   -    DELETED FROM OUTDAT
```

```
C       SPECNO - INCORPORATED INTO VARIABLE SPCNO
C       MAXAMP - SHIFTED TO SAVE EXECUTION TIME
C       TOTION.TICRAT - UNMODIFIED
C                                      WM F HARGROVE
C                                      12/6/78
C      EQUIVALENCE(EXPID.LIBBLK(84))
C      EQUIVALENCE(PDATE.LIBBLK(92))
C      EQUIVALENCE (AMP.LIBBLK(257))
C      EQUIVALENCE(LOMASX.LIBBLK(75))
C      EQUIVALENCE(HIMASX.LIBBLK(74))
C      EQUIVALENCE(SPECNO.LIBBLK(78))
C      EQUIVALENCE(MAXAMP.LIBBLK(77))
C      EQUIVALENCE(TOTION.LIBBLK(1))
C      EQUIVALENCE(LSTCNX.LIBBLK(70))
C      EQUIVALENCE(GCTIMX.LIBBLK(90))
C      EQUIVALENCE(INTEGX.LIBBLK(89))
C      EQUIVALENCE(ABMAX.LIBBLK(49))
C      EQUIVALENCE(TICRAT.LIBBLK(72))
       DATA HOW/1H /
C       **********************HERE WE GO***************************************
C       TAPCNT 1000 IF WE ARE GOING TO WRITE THE BIT MAP
       IF ( TAPCNT .EQ. 1000 ) GO TO 12
C THIS MOD INSERTED  TO WEED OUT ELLUTANTS WHICH HAVE NO
C SUITABLE MODEL.....
C                         RTI WFH 11/21/78
       IF(IGMASS.GT.0)GO TO 91
       IF(IDEBUG.NE.0)WRITE(IPTR2.3334)
 3334  FORMAT(1X.' SPECTRUM REJECTED FOR LACK OF SUITABLE MODEL')
       RETURN
C
   91  ISWEAK = IGMASS - 39
       IF ( NUSPEC(ISWEAK) .GT. 0 ) GO TO 93
       IF ( IDEBUG .EQ. 0 ) GO TO 92
       INEWM=IGMASS-MAD
C      CALL TRACE(IGMASS.MAD)
C THIS WRITE STMT HAS BEEN MADE CONDITIONAL TO SAVE PAPER
C                                             WFH RTI 9/31/78
       IF(IDEBUG.EQ.0)GO TO 92
       WRITE(IPTR2.3333) INEWM
92     RETURN
93     CONTINUE
C      CALL ZERO(LIBBLK.1024)
C SEE COMMENT AT TOP
C                    WFH - RTI 12/6/78
C      DO 30 I=1.1024
C30    LIBBLK(I)=0
       TOTION=0.0
C      CALL TRACE(NMASS)
       DO 4  I = 1.NMASS
4      IF  ( NUSPEC(I) .LT. 0 ) NUSPEC(I) = 0
       SATSUM = 0
       DO 5  I = 1.NMASS  .
       NUZEN = NUSPEC(I)
       IF ( NUZEN .GE. SATVAL ) SATSUM = SATSUM + NUZEN
5      TOTION = TOTION + NUZEN
C       REJECT SPECTRUM IF TOTION IS LESS THAN MINIMUM TIC
       IF ( TOTION  .LT.  MINTIC ) RETURN
C       TICRAT IS CLEANUP RATIO
```

```
      SIGMA7 = SIGMA7 - QSATS * SATVAL + SATSUM
C        SINCE COMPUTED SATURATED VALUES MAY CAUSE THE CLEANED UP
C        TIC TO EXCEED THE RAW, WE SUBSTITUTE CALCULATED VALUES
C        FOR MAXIMUM VALUES. THIS IS  A LIMIT OF OUR 16 BIT INTEGER.
C        AND 24570 SATURATED DEFAULT VALUE IN OUR GC COLLECTION
C        PHASE.
      IF ( TOTION .GE. SIGMA7 ) TOTION = SIGMA7 - 1
      TICRAT=TOTION*100.0/SIGMA7
      IF ( IDEBUG .EQ. 0 ) GO TO 6
      WRITE(IPTR2.1000) TOTION.SIGMA7.TICRAT.SATSUM
6     OUTSLT=0
      LSTCNT=LSTCNT+1
C     CALL MOVE(EXPNAM.EXPID.5)
C SEE COMMENT AT TOP
C                     WFH - RTI 12/6/78
C     DO 39 I=1.5
C39   EXPID(I)=EXPNAM(I)
C     CALL MOVE(NAME.PNAME.32)
C THIS LOOP COMENTED OUT TO SUIT WATFIV AND BECAUSE IT SERVES
C NO APPARENT PURPOSE...
C                          WFH AT RTI 4/25/78
C     DO 32 I=1.32
C 32  PNAME(I)=NAME(I)
C SEE COMMENT AT TOP
C                       WFH - RTI 12/6/78
C     INTEGX=INTEG
C     CALL MOVE(RDATE.PDATE.5)
C   SEE COMMENT AT TOP
C                      WFH - RTI 12/6/78
C     DO 31 I=1.5
C31   PDATE(I)=RDATE(I)
C     LOMASX=LOMASS
C     HIMASX=HIMASS
C     LSTCNX=LSTCNT
C     SPECNO=SPCNO
C     CALL TRACE(NMASS)
C    MAXAMP DETERMINATION SHIFTED ...
C                                WFH - RTI 12/6/78
      IF ( IPFLAG .EQ. 0 ) GO TO 12
      MAXAMP=0
      DO 10 I=1.NMASS
   10 IF(MAXAMP.LT.NUSPEC(I)) MAXAMP=NUSPEC(I)
C      COMPUTE ABSOLUTE MAXIMUM AMPLITUDE.
      ABMAX = MAXAMP/QNORM
      CALL KAPUT( SPCNO.TOTION.TICRAT.ABMAX )
C      THE ABOVE PRINTS OUT THE "CHEMISTS" MINIMUM STUFF.
11    CONTINUE
C     CALL SETMAS(LIBMAP.LSTCNT)
C ...     CALL PUTFIL(OUTSLT.LIBBLK.LSTCNT)
C14      CALL CHKFIL(OUTSLT.FLAG)
C.....     IF(FLAG) 13.14.15
C..15       CONTINUE
C     WRITE(IPTR3.34) (LIBBLK(I).I=1.128)
C  34 FORMAT(3(/32I4)/13I4.10A4.9I4/)
      IZ=0
      DO 81 I=1.411
      IF(NUSPEC(I).LE.0)GO TO 81
      IZ=IZ+1
      JM(IZ)=I+39-MAD
C      MM(IZ)=(I+39-MAD)*10
      RMM(IZ)=I+39-MAD
      JSP(IZ)=NUSPEC(I)
C      MSP(IZ)=NUSPEC(I)
      RMSP(IZ)=NUSPEC(I)
81    CONTINUE
C THESE WRITE STMTS HAVE BEEN MADE CONDITIONAL FOR CONSISTENCY
C                                            WFH RTI 9/31/78
      IF(IDEBUG.EQ.0)GO TO 85
      WRITE(IPTR2.80)SPCNO.IZ.(NNAME(K).K=1.15)
80    FORMAT(4X.'SPECTRA = '.I5.'  NO. OF PEAKS = '.I4.2X.15A4.' CLEANE
     *D.'/)
      WRITE(IPTR2.83)(JM(K).JSP(K).K=1.IZ)
83    FORMAT(46(18I6/)/)
85    CALL XTREM(JSP.IZ.MIN.MINI.MAX.MAXI)
      BASE=MAX
C     CALL TRACE(IZ)
      DO 60 I=1.IZ
60    RMSP(I)=(RMSP(I)/BASE)*100.
      RMSP(MAXI)=99.99
C     WRITE(8.600)SPECNO.IZ.(NNAME(I).I=1.15).CLASS.SUBCL.HOW.IMOLWT
C THE FOLLOWING TWO WRITE STATEMENTS WERE CONENTED OUT BECAUSE
C THEY SEEM TO YEILD NOTHING INTERESTING AND ALSO BECAUSE NO
C CARRIAGE CONTROL FOR THE PRINTER HAS BEEN INCLUDED
C                                           WM.F.HARGROVE - RTI
C     WRITE(IPTR2.600)SPECNO.IZ.(NNAME(I).I=1.15).CLASS.SUBCL.HOW.IMOLWT
C600  FORMAT(I5.I3.15A4.9A1.I3)
C
C HERE IS THE MAIN OUTPUT STATEMENT WRITING THE CLEANUPED
C UP SPECTRUM TO A DISK DATASET
C THE CHANGE TO A SUBROUTINE CALL HAS BEEN MADE WITH THE
C ONSET OF MULTIPLE RUN PROCESSING AND A LINEAR SYSTEM.
C                                         WM F HARGROVE
C                                         11/29/78
C     WRITE(9) SPCNO.NUSPEC
      CALL WRTSPC(SPCNO.NUSPEC.QDUBLT.TICRAT.IZ.TOTION)
      N=(IZ+7)/8
      IF=1
      IT=8
      DO 62 I=1.N
C THIS MOD INSERTED TO BLOT OUT APPARENTLY USELESS WRITE STMT.
C                             WFH - 5/31/78 - RTI
C     WRITE(8.601)(RMM(K).RMSP(K).K=IF.IT)
C ANOTHER WRITE STMT MADE CONDITIONAL FOR CONSISTENCY'S SAKE
C                                            WFH RTI 9/31/78
      IF(IDEBUG.EQ.0)GO TO 63
      WRITE(IPTR2.602)(RMM(K).RMSP(K).K=IF.IT)
   63 CONTINUE
      IF=IT+1
62    IT=IF+7
C601  FORMAT(8(F5.1.F5.2)).
602   FORMAT(1X.8(F5.1.F5.2))
      DO 61 I=1.411
      RMM(I)=0.0
61    RMSP(I)=0.0
C        BASM=MM(MAXI)/10
C        BASE=MAX
```

```
C          DO 50 I=1.12
C          HD1=JSP(I)
C     50   MSP(I)=(HD1/BASE)*10000
C....
C.....
C       WRITE(8)MW.I1.I2.BASM.NNAME.SOURCE.FORM.(MM(K).MSP(K).K=1.I2)
12    CONTINUE
C      DON'T WRITE BITMAP UNTIL LAST RECORD.
C     RETURN
C      HERE FOR WRITING BITMAP TO LIST.MAS
C..12        CALL ZERO(LIBBLK.1024)
C  12   DO 33 I=1.1024
C  33   LIRBLK(I)=0
C     OUTSLT = 0
C     LSTCNX=LSTCNT
C      READ(7.20)LIBBLK
C   20  FORMAT(I8)
C     CALL GETFIL(OUTSLT.LIBBLK.0)
C..20        CALL CHKFIL(OUTSLT.FLAG)
C     IF(FLAG) 13.20.21
C..21        CONTINUE
C..       CALL MOVE(LIBMAP.LIBBLK.17)
C         DO 35 I=1.17
C      35 LIBBLK(I)=LIBMAP(I)
C         WRITE(IPTR3.36)LIBBLK
C      36 FORMAT(I8)
C..       CALL PUTFIL(OUTSLT.LIBBLK.0)
C..16        CALL CHKFIL(OUTSLT.FLAG)
C..       IF(FLAG) 13.16.17
C..13        CALL PUTNUM('FILE ERROR FLAG = '.FLAG)
C..       CALL PUTSTR('BOMBED WRITTING TO LIST.MAS IN OUTDAT')
C..       STOP
C..17        CALL CLSFIL(OUTSLT)
      RETURN
1000  FORMAT(/.' TOTAL ION CURRENT ='.F8.0./.' UNC SIG ='.E15.8.
     -/' TICRAT ='.I3/' SAT PEAK SUM ='.F12.0//)
3333  FORMAT(/.' SPECTRUM REJECTED BECAUSE MASS '.I3.'.THE '.
     1' MODEL PEAK. HAS 0 CONTRIBUTION        '/)
      END
      SUBROUTINE PARDIR
C
C       THIS COMPUTER PROGRAM WAS DEVELOPED WITH THE FUNDING SUPPORT
C       FROM THE NATIONAL INSTITUTES OF HEALTH (GRANTS RR-612 AND
C       GM-20832) AND THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
C       (GRANT NGR-05-020-632).
C FURTHER NOTE:
C THIS VERSION OF PARDIR HAS BEEN COMPLETELY REWRITTEN IN ORDER TO ALLOW
C FOR PROCESSING OF MULTIPLE RUNS. THIS VERSION OF PARDIR IS ALSO USED
C IN THE VERSIN OF CLEANUP SUPPLIED TO EPA BY RTI.THE FUNCTION OF
C PARDIR IS TO READ IN CONTROL PARAMETERS FROM A TERMINAL OR
C CONTROL CARDS THAT ARE CONSTANT ACROSS DIFFERENT RUNS. THE
C MEANING OF THE VARIABLES READ HEREIN SHOULD BE SELF-DOCUMENTING
C FROM THE ECHO MESSAGES.
C
C                                               WM.F.HARGROVE
C                                               RTI 11/22/78
C
      COMMON /UNITS/ IRDR.IPTR1.IPTR2.IPTR3.IIN.IOUT
      INTEGER IRDR.IPTR1.IPTR2.IPTR3.IIN.IOUT
      COMMON /TOMS/ITOM
      COMMON /PRNTFG/IPFLAG
      COMMON /LFTRIT/ILEFT.IRIGHT
      COMMON /CRATMX/RATMXX
      COMMON /DUBBUG/IDEBUG
      COMMON /PARAMS/MINTIC.MINNY.NTHIRD.RATM
      COMMON /AVG/ NAVG.HAVG
      INTEGER ITOM.IPFLAG.ILEFT.IRIGHT.IDEBUG.NTHIRD
      REAL RATMXX.RATM
      INTEGER NAVG.HAVG
      INTEGER MINNY.MINTIC
C PROMPT AND READ CONTROL PARMS - LIFTED FROM OLD CLEANUP SOURCE
C AN ADDITIONAL PARAMETER WAS ADDED TO AVERAGE N INPUT
C SPECTRA IF IT BECAME NECESSARY - AS WITH THE INCOS DATA
C                                             RTI - WFH 1/14/79
      WRITE(IPTR1.3240)
 3240 FORMAT(1X.'ENTER CLEANUP PRINT FLAG(IN I1):')
      READ(IRDR.3241) IPFLAG
 3241 FORMAT(I1)
      WRITE(IPTR1.3242) IPFLAG
 3242 FORMAT(1X.'IPFLAG=',I1)
C
C       IPFLAG IS PRINT OPTION FLAG FOR CLEANUP
C       INITIALIZE DEFAULTS FOR VARIOUS PARAMETERS IN COMMON
      WRITE(IPTR1.3100)
 3100 FORMAT(1X.'ENTER DEBUGGING SWITCH IN I1 (0=N.1=Y):')
      READ(IRDR.3101) IDEBUG
 3101 FORMAT(I1)
      WRITE(IPTR1.3103) IDEBUG
 3103 FORMAT(1X.'IDEBUG=',I1)
C
      WRITE(IPTR1.123)
  123 FORMAT(1X.'ENTER # OF INPUT SPECTRA TO AVERAGE IN I2:')
      READ(IRDR.121) NAVG
  121 FORMAT(I2)
      WRITE(IPTR1.122) NAVG
  122 FORMAT(1X.'NAVG=',I2)
      HAVG=NAVG/2+1
      WRITE(IPTR1.3120)
 3120 FORMAT(1X.'MINIMUM THIRDS BETWEEN ELLUANTS IN I1:')
      READ(IRDR.3121) NTHIRD
 3121 FORMAT(I1)
      WRITE(IPTR1.3123) NTHIRD
 3123 FORMAT(1X.'THIRDS BETWEEN ELLUANTS=',I1)
C
      WRITE(IPTR1.3130)
 3130 FORMAT(1X.'ENTER DETECTION WINDOW IN I4:')
      READ(IRDR.3131) ITOM
 3131 FORMAT(I4)
      WRITE(IPTR1.3132) ITOM
 3132 FORMAT(1X.'DETECTION WINDOW=',I4)
C
      WRITE(IPTR1.3140)
 3140 FORMAT(1X.'ENTER MINIMUM FOR LT 5 PEAKS IN I6:')
      READ(IRDR.3141) MINNY
 3141 FORMAT(I6)
      WRITE(IPTR1.3:43) MINNY
 3143 FORMAT(1X.'MINNY=',I6)
```

```
C
      WRITE(IPTR1.3150)
 3150 FORMAT(1X.'ENTER MINIMUM FOR MORE THAN 4 PEAKS IN I6:')
      READ(IRDR.3151) MINTIC
 3151 FORMAT(I6)
      WRITE(IPTR1.3152) MINTIC
 3152 FORMAT(1X.'MINTIC=',I6)
C
      WRITE(IPTR1.3160)
 3160 FORMAT(1X.'ENTER MINIMUM QUALIFYING RATE IN F8.4:')
      READ(IRDR.3161) RATM
 3161 FORMAT(F8.4)
      WRITE(IPTR1.3162) RATM
 3162 FORMAT(1X.'RATM=',F8.4)
C
      WRITE(IPTR1.3170)
 3170 FORMAT(1X.'ENTER ERROR RATIO IN F8.4:')
      READ(IRDR.3171) RATMXX
 3171 FORMAT(F8.4)
      WRITE(IPTR1.3172) RATMXX
 3172 FORMAT(1X.'RATMXX=',F8.4)
C
      WRITE(IPTR1.3180)
 3180 FORMAT(1X.'ENTER POINTS TO LEFT OF MODE IN I3:')
      READ(IRDR.3181) ILEFT
 3181 FORMAT(I3)
      WRITE(IPTR1.3182) ILEFT
 3182 FORMAT(1X.'ILEFT='.I3)
C
      WRITE(IPTR1.3190)
 3190 FORMAT(1X.'ENTER POINTS TO RIGHT OF MODE IN I3:')
      READ(IRDR.3191) IRIGHT
 3191 FORMAT(I3)
      WRITE(IPTR1.3192) IRIGHT
 3192 FORMAT(1X.'IRIGHT= '.I3)
      RETURN
      END
      SUBROUTINE SATLSQ (NS.M.RPKHGT.G.SPEC.SATVAL.GPOS.OVLAP.HPOS.PKHIS
     *T)
C       THIS COMPUTER PROGRAM WAS DEVELOPED WITH THE FUNDING SUPPORT
C       FROM THE NATIONAL INSTITUTES OF HEALTH (GRANTS RR-612 AND
C       GM-20832) AND THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
C       (GRANT NGR-05-020-632).
C
C
C       THIS ROUTINE COMPUTES SATURATED VALUES IN THE
C       FOLLOWING WAY:
C
C           1. REMOVE LOCAL BACKGROUND FROM SATURATED PROFILE.
C
C           2. COMPUTE THE RATIOS OF PEAK(I)/MODEL(I)
C              FOR NON-SATURATED POINTS IN THE PEAK.
C
C            3. GET THE AVERAGE RATIO. AND MULTIPLY IT TIMES
C              THE MODEL MODE. THIS IS THE RETURNED SATURATED
C              VALUE.
      REAL     LOCLBG.GG.PKHIST(20)
      INTEGER SPEC(14.411).GPOS.OVLAP.HPOS
      INTEGER MINHGT.SATLST.SATINX
      INTEGER IJ.IK.IL.LSTOP.RSTART.STOPF.RCOUNT
      DIMENSION G(9).Y(13).RATIOS(13)
      COMMON /CSATLS/SATLST(25).SATINX
      COMMON /MODES/SGMODE(3).SHMODE(3).GMODE.HMODE
      COMMON /QUIXOT/IQUIT
C      I2=NS-12-14*((NS-13)/14)
C      I3=NS-11-14*((NS-12)/14)
C      I4=NS-10-14*((NS-11)/14)
C      I5=NS-9-14*((NS-10)/14)
C      I6=NS-8-14*((NS-9)/14)
C      I7=NS-7-14*((NS-8)/14)
C      I8=NS-6-14*((NS-7)/14)
C      I9=NS-5-14*((NS-6)/14)
C      I10=NS-4-14*((NS-5)/14)
C      I11=NS-3-14*((NS-4)/14)
C      I12=NS-2-14*((NS-3)/14)
C      I13=NS-1-14*((NS-2)/14)
      IK = 13
      DO 5    I = 1.12
      IL = IK - 1
      IJ = NS - IL - 14 * ((NS - IK)/14)
      Y(I) = SPEC(IJ.M)
      IK = IK - 1
5     CONTINUE
C      SATVAL WILL BE DEFAULT IN CASE OF SUPER SATURATION
      RPKHGT = -SATVAL
      MINHGT = .05 * SATVAL
C      MINHGT IS MINIMUM ACCEPTABLE HEIGHT IN MODEL FOR RATIOING.
      ILFT = 0
      ILEFT = 0
      IRIGHT = 0
      NPTS = 12
      NSATS = 0
      DO  50 I = 1.NPTS
      IF (Y(I).GE.SATVAL) GO TO  10
      IF (ILFT.EQ.1) GO TO  20
C
C      COUNT POINTS TO LEFT OF FIRST SATURATED POINT
C
      ILEFT = ILEFT + 1
      GO TO  50
20    CONTINUE
C      COUNT POINTS TO RIGHT OF LAST SATURATED POINT
      IRIGHT = IRIGHT + 1
      GO TO  50
C      FOUND FIRST SATURATED POINT
10    ILFT = 1
      NSATS = NSATS + 1
C      NSATS COUNTS SATURATED VALUES.
50    CONTINUE
C      ILEFT = NUMBER OF POINTS TO LEFT OF FIRST SATURATED POINT.
C      IRIGHT = NUMBER OF POINTS TO RIGHT OF LAST SATURATED POINT.
C      FIND MIDDLE POINT OF SATURATED VALUES AND D MAKE SURE
C      7 IS IN +MID-1.MID+10.NOTE:THE INTERVAL IS CLOSED.
      IQUIT = 0
      IFIRST = ILEFT + 1
      ILAST = ILEFT + NSATS
```

```
C
C        LLLLLSSSSRRR
C
C        IFIRST ILAST
C
C        RMID IS MIDDLE POINT OF SATURATED VALUES. AND REAL.
        RMID = ( IFIRST + ILAST )/2.0
        RWIND = 7.0 - RMID
C        NOTE: 7 IS INDEX OF CURRENT SPECTRUM.
C         7 MUST BE WITHIN (RMID-1.RMID+1).
C
C        ACTUALLY. RMID MUST BE IN +6.80 FOR SATURATED
C        VALUE TO BE PROCESSED.
        IF ( RWIND .LT. 0 ) RWIND = -RWIND
        IF ( RWIND .LE. 1.0 ) GO TO 60
        IQUIT = 1
        RETURN
60      CONTINUE
C        PRINT SATURATED MASS ENCOUNTERED STUFF
        MASS = M + 39
C        SEE IF MASS BELONGS TO PREDESSOR.
        DO 62 I = 1.SATINX
        IF ( SATLST(I) .NE. MASS ) GO TO 62
        SATLST(I) = 0
        IQUIT = 2
        RETURN
62      CONTINUE
C        DETERMINE IF THIS MASS BELONGS TO AN AJACENT SPECTRUM
        IF ( OVLAP .EQ. 0 )  GO TO 70
        NXTSPC = 6 + ( HPOS-1 )/3
C        SEE IF SATVALS OCCUR IN NEXT SPECTRUM CONTINUOUSLY
        IF ( ILAST .LT. NXTSPC ) GO TO 70
        RDIFF = NXTSPC - RMID
C        SEE IF RMID IS CLOSE TO NEXT SPECTRUM.
        IF ( RDIFF .GT. 1 ) GO TO 70
C        GIVE UNDECIDED MASS TO SPECTRUM WITH BIGGEST TIC.
        IF ( PKHIST(GPOS) .GE. PKHIST(HPOS) ) GO TO 65
        IQUIT = 1
        RETURN
65      CONTINUE
C        FLAG FOR NEXT GUY TO AVOID.
        SATLST(SATINX) = MASS
        SATINX = SATINX + 1
        IF ( SATINX .EQ. 26 ) SATINX = 25
70      CONTINUE
C THE LAST THREE ARGUMENTS Y.Y.Y) HAVE BEEN CHANGED TO
C SATVAL.SATVAL.SATVAL) TO SUIT WATFIV...THEY HAVE NO
C FUNCTION WITHIN SPEPRT.
C                                        WFH RTI 4/30/78
        CALL SPEPRT(1.MASS.Y.SATVAL.SATVAL.SATVAL)
C THE LAST TWO ARGUMENTS Y.Y) HAVE ALSO BEEN CHANGED TO SATVAL.SATVAL)
C TO SUIT WATFIV ... THEY HAVE NO APPARENT FUNCTION INSIDE SPEPRT
C                                                WFH RTI 4/30/78
C        PRINT DEFAULT SATURATED VALUE
        CALL SPEPRT(2.NPTS.Y.SATVAL.SATVAL.SATVAL)
C        ********  NOW BEGIN OUR SATURATION STUFF **********
        LSTOP = ILEFT
        RSTART = NPTS - IRIGHT + 1
```

```
      LOCLBG = Y(1)
C       SUBTRACT LEFT POINT FROM ALL VALUES AS AN APPROXIMATION
C       OF BACKGROUND.THIS IS TO BE SYMMETRIC WITH MODEL PEAK.
      STOPF = 0
      IF ( LSTOP .NE. 0) GO TO 90
      STOPF = 1
      GO TO 105
90    CONTINUE
      DO 100 I = 1,LSTOP
100   Y(I) = Y(I) - LOCLBG
105   CONTINUE
      IF (STOPF .EQ. 1  .AND.  RSTART .GT. NPTS) GO TO 900
C       IF WE GO TO 900 IT MEANS ALL VALUES ARE SATURATED
      IF (RSTART .GT. NPTS) GO TO 120
      DO 110 I = RSTART.NPTS
110   Y(I) = Y(I) - LOCLBG
120   CONTINUE
      RATO = 0
      RCOUNT = 0
      IF (LSTOP .LT. 3) GO TO 300
      DO 200 I = 3.LSTOP
      GG = G(I-2)
      IF( GG .LT. MINHGT ) GO TO 200
C    NOTE: MINGHT > 0
      RATIOS(I) = Y(I)/GG
      RATO = RATO + RATIOS(I)
      RCOUNT = RCOUNT + 1
200   CONTINUE
300   CONTINUE
      IF (RSTART .GT. 11) GO TO 350
      DO 325 I = RSTART,11
      GG = G(I-2)
      IF ( GG .LT. MINHGT ) GO TO 325
      RATIOS(I) = Y(I)/GG
      RATO = RATO + RATIOS(I)
      RCOUNT = RCOUNT + 1
325   CONTINUE
350   CONTINUE
      IF(RCOUNT .LE. 1) GO TO 900
C       ABOVE TAKES CARE OF SUPER SATURATION
      RATO = RATO/RCOUNT
      IF(RATO .LE. 0) GO TO 900
C       THE ABOVE TAKES CARE OF A VERY FUNNY LOCAL BACKGROUND
      RPKHGT = RATO * G(5)
C       THE ABOVE TAKES CARE OF SATURATED PEAKS WITH TO HIGH
C       OF BACKGROUND APPROXIMATIONS.
      CALL SPEPRT(3.I.RPKHGT.LOCLBG,RATO,G(5))
      RETURN
900   CONTINUE
      CALL SPEPRT(4,I.RPKHGT.RATO,RATO,RATO)
      RETURN
      END
C
C THIS SUBROUTINE WAS INCORPORATED TO ACCOUNT FOR THE POSSIBILITY
C OF DIFFERENT SCAN FUNCTIONS ON THE FINNIGAN QUADROPOLE
C SPECTROMETER AND TO ALLOW FOR MAGNETIC INSTRUMENTS IN THE FUTURE
C                                              WFH RTI 9/31/78
C
```

```
      REAL FUNCTION SCAN(MASS)
      INTEGER*4 MASS
      REAL*4 HIMASS/450./,LOMASS/40./
C
      RMASS=MASS
      SCAN=(RMASS-LOMASS)/(HIMASS-LOMASS)
      RETURN
      END
      SUBROUTINE SETBUG(NNAME.EOF)
C
      INTEGER CLNFLG
C      THIS COMPUTER PROGRAM WAS DEVELOPED WITH THE FUNDING SUPPORT
C      FROM THE NATIONAL INSTITUTES OF HEALTH (GRANTS RR-612 AND
C      GM-20832) AND THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
C      (GRANT NGR-05-020-632).
C
C FURTHER NOTE:
C THIS VERSION OF SUBROUTINE SETBUG HAS REWRITTEN IN ORDER TO PERMIT
C MULTIPLE RUN PROCESSING AND TO FIT INTO THE CONTEXT OF THE EPA AND
C RTI ACQUISITION SYSTEMS. ITS FUNCTION IS TO READ AND FILL THE SPECTRUM
C BUFFER WITH THE FIRST 13 SPECTRA FROM EACH RUN.
      COMMON PKHIST(20).NPHIST(20)
      COMMON YM5.YM4.YM3.YM2.YM1.Y0.YP1.YP2.YP3.YP4.YP5
      COMMON GM4.GM3.GM2.GM1.G0.GP1.GP2.GP3.GP4
      COMMON SG.SG2.SIG.SI.SI2.AN.AA.BB.CC.DD
      COMMON NAME(32).RDATE(5).BITMAP(64).GCTIME
      COMMON LIBMAP(17).INTEG.LOMASS.HIMASS.POINTS.TAPCNT
      COMMON MAXSP.LSTCNT.DEVICE(5).SPCNO
      COMMON SPEC(14.411).IND(3.411).NUSPEC(411).NMASS
      COMMON RMAX(3).IR(3).BK(3).NM(3).NTM(3).GPEAK(3.11)
      COMMON SPARE1.SPARE2.SPARE3.SPARE4.SPARE5.SPARE6.SPARE7
      COMMON RATMIN.RATEMX.SIGY0.SIGNXT.SIGLST
      COMMON ISVN8.ISVN9.ISVP8.ISVP9.SIGMA7.SATVAL.SATMAX
      COMMON MODE1.MODE2.NLWCUT.NUPCUT.SATPKS(25).NSAT.EXPNAM(5)
      REAL SATVAL
      INTEGER AVGSPC(411).NAVG.HAVG
      INTEGER EOF.LIBBLK.RDATE.LOMASS.HIMASS.ADS.POINTS,
     1SPCNO.SPEC.NNAME(18)
      COMMON /SCR/ LIBBLK(1024)
      COMMON /IADS/ ADS
      COMMON /IMAD/ MAD
      COMMON /AVG/ NAVG.HAVG
C READ IN RUN HEADER AND RETURN IF END OF FILE
      CALL OPNRUN(LIBBLK.EOF)
      IF(EOF.NE.0)RETURN
C HOUSE CLEANING FOR HISTORICAL VARIABLES STILL LEFT IN CLEANUP
      NMASS=411
      DO 10 I=1.2
   10 RDATE(I)=LIBBLK(I+27)
C SETTING OF HIGH MASS. LOMASS AND OF • OF SPECTRA IN RUN
      MAXSP=LIBBLK(38)
      LOMASS=40
      HIMASS=450
      ADS=LIBBLK(36)
      MAD=40-LOMASS
      POINTS=40
C BY SERENDIPIDY NNAME CAN BE MODIFIED TO CONTAIN THE RUN TITLE AND
C EXPNAM TO CONTAIN THE INSTRUMENT TYPE...
```

```
      DO 20 I=1.10
   20 NNAME(I)=LIBBLK(12+I)
      DO 30 I=1.5
   30 EXPNAM(I)=LIBBLK(7+I)
C THE INTERNALS OF SETBUG HAVE BEEN MODIFIED SO AS TO
C INCLUDE AN AVERAGING OVER THE INPUT SPECTRA
C                                         RTI - WFH 1/13/79
      DO 50 I=1.13
C INITIALIZE MAIN SPECTRUM ARRAY IN CORE FOR ADDITION
      DO 31 IK=1.411
   31 SPEC(I.IK)=0
C READ IN NAVG SPECTRA FOR THE CURRENT CORE POSITION
      DO 40 IJ=1.NAVG
      IQUPD=0
C IQUPD ENSURES THAT THE QUE IN THE I/O ROUTINES IS
C UPDATED ONLY ONCE
C I14 TELLS WHICH QUEUE LOCATION TO UPDATE
      IF(IJ.EQ.HAVG)IQUPD=1
      CALL RDSPC(ISCNO.AVGSPC.I.IQUPD.EOF)
      IQUPD=0
      IF(EOF.EQ.0)GO TO 34
C EOF HAS BEEN ENCOUNTERED BEFORE ALL 13 CORE POSTIONS WERE FILLED
      WRITE(IPTR1.60)
   60 FORMAT('-'.'*** UNABLE TO FILL 13 CORE POSITIONS - EOF ***')
      STOP
   34 CONTINUE
C NOW INCREMENT MAIN ARRAY IN CORE
      DO 32 IK=1.411
   32 SPEC(I.IK)=SPEC(I.IK)+AVGSPC(IK)
   40 CONTINUE
C NOW DIVIDE BY NAVG IF NECESSARY
      IF(NAVG.EQ.1)GO TO 47
      DO 46 IK=1.411
   46 SPEC(I.IK)=SPEC(I.IK)/NAVG
   47 CONTINUE
   50 CONTINUE
C NOW SET INTERNAL SPECTRUM NUMBER
      SPCNO=ISCNO-6*NAVG-HAVG
      RETURN
      END
      SUBROUTINE SETGHY(NS.M.GNU.HNU.GPOS.HPOS.ISW.GMX)
C
C      *** THIS ROUTINE ACTUALLY DOES THE SETUP AND
C      *** INTERPOLATION PRIOR TO DOUBLET RESOLUTION
C      *** IT CALLS THE DOUBLET RESOLVER.
C
      INTEGER NAME.RDATE.BITMAP.GCTIME.LIBMAP.INTEG.LOMASS
      INTEGER HIMASS.POINTS.TAPCNT.MAXSP.LSTCNT.DEVICE.SPCNO
      INTEGER SPEC.IND.IR.BK.NM.NTM
      INTEGER GPEAK.SPARE1.SPARE2.SPARE3.SPARE4.SPARE7
      REAL    SPARE5.SPARE6
      INTEGER GPOS.HPOS
      INTEGER NUSPEC.NMASS.OVLAP.ISW
      INTEGER STORE
      REAL    ISVP8.ISVP9
      REAL    DMX
      REAL    GMX
C
```

```
      DIMENSION GNU(18).HNU(18).Y(18)
      COMMON PKHIST(20).NPHIST(20)
      COMMON YM5.YM4.YM3.YM2.YM1.Y0.YP1.YP2.YP3.YP4.YP5
      COMMON GM4.GM3.GM2.GM1.G0.GP1.GP2.GP3.GP4
      COMMON SG.SG2.SIG.SI.SI2.AN.AA.BB.CC.DD
      COMMON NAME(32).RDATE(5).BITMAP(64).GCTIME
      COMMON LIBMAP(17).INTEG.LOMASS.HIMASS.POINTS.TAPCNT
      COMMON MAXSP.LSTCNT.DEVICE(5).SPCNO
      COMMON SPEC(14.411).IND(3.411).NUSPEC(411).NMASS
      COMMON RMAX(3).IR(3).BK(3).NM(3).NTM(3).GPEAK(3.11)
      COMMON SPARE1.SPARE2.SPARE3.SPARE4.SPARE5.SPARE6.SPARE7
      COMMON RATMIN.RATEMX.SIGY0.SIGNXT.SIGLST
      COMMON ISVN8.ISVN9.ISVP8.ISVP9.SIGMA7.SATVAL.SATMAX
      COMMON MODE1.MODE2.NLWCUT.NUPCUT.SATPKS(25).NSAT.EXPNAM(5)
      COMMON /MODES/SGMODE(3).SHMODE(3).GMODE.HMODE
      COMMON /CRATMX/RATMXX
      INTEGER BEDROK
      DATA BEDROK/40/
C
C
C      DO ROTARY BUFFER INDEXING OF SPEC(14.NUMMASSES)
      J = 13
      DO  1000  I=1.13
      K = J - 1
      IJ = NS - K - 14 * ((NS - J)/14)
      Y(I) = SPEC(IJ.M)
      J = J - 1
1000  CONTINUE
C      DO NOT LET ANY PEAKS WITH SATURATED VALUES IN THEIR
C      PROFILES PASS THROUGH DOUBLET RESOLVER.
      DO 50  I= 1.11
      IF ( Y(I) .GE. SATVAL ) GO TO 5
50    CONTINUE
C
C    WE TRY TO ELIMINATE BACKGROUND PEAKS FROM PASSING THROUGH
C    DOUBLET FILTER: SO. LET K1 = .0125 * Y(7). AND K2 = .0125 * G(7).
C    THEN: WHERE M = MAX↑K1.K20. LET D = MIN↑M.700.AND D = MAX↑D.300
C          (WE USE AN ABSOLUTE MAX OF 70 SO AS NOT TO MISS HIGH MASS
C           PEAKS WITH SMALL CONTRIBUTIONS WHEN THE BASE PEAK IS
C           LARGE.AND A MIN OF 30 FOR LOW AMPLITUDE DRIBBLE.)
C
C     ------------------------------------------  ←  Y(7) + E
C          .                          .    .        E
C     .        .      * ←  Y(7)              .     -
C      .                 .                   E
C     ------------------------------------------  ←  Y(7) - E
C
C      IF THE FRAGMENTOGRAM STAYS IN THE ABOVE ENVELOPE FROM Y(3) TO
C       TO Y(11). THEN WE CONSIDER IT A BACKGROUND PEAK.
C
      D = .0125
      K1 = D * Y(7)
C      WE MUST RECALL THAT G(7) HAS BEEN NORMALIZED IN GETHGT
      K2 = D * GNU(7)/GMX
      DMX = K1
      IF ( DMX .LT . K2 ) DMX = K2
C      DMX IS MAX↑K1.K20
      IF( DMX .GT. 70) DMX = 70
```

```
C        DMX = MIN+DMX.700
        IF(DMX .LT. BEDROK ) DMX  = BEDROK
C        DMX = MAX+BEDROK.DMX0
       E1 = Y(7) + DMX
       E2 = Y(7) - DMX
       DO 100 I = 3,11
       IF (Y(I) .GT. E1   ,OR.   Y(I) .LT. E2) GO TO 200
100    CONTINUE
       GHGT = 0
       RETURN
200    CONTINUE
C
C       WE NOW TRY TO ELIMINATE LOCALLY "NEARLY LINEAR"
C       CONTRIBUTIONS BY FITTING A LINE TO THE LOCAL FRAGMENTOGRAM
C       IN A LEAST SQUARES WAY, AND EXAMINING
C
C           RATIO =  (ROOT MEAN SQUARE ERROR)/(AVERAGE VALUE OF Y(I))
C
C
C       THIS TEST HAS  BEEN VERY SUCESSFUL IN ELIMINATING COLUMN
C       BLEED .
      CALL BLEED(Y(5).RATIO)
C THIS MOD INSERTED ON A TEMPORARY BASIS TO GATHER STATISTICS
C ON ERROR RATIO - CODE 3
C                                         WFH RTI 9/31/78
C      ITYP=3
C      ISPC=SPCNO
C      WRITE(20.6250) ITYP.ISPC.RATIO.RATIO
C6250 FORMAT(2A4.20X.2A4)
       IF (RATIO .GT. RATMXX) GO TO 290
      GHGT = 0
      RETURN
290   CONTINUE
C    ALSO. IF Y(I) IS NON INCREASING FOR I=4.9. IE.Y(I) .GE. Y(I+1).
C         THEN DON'T CONSIDER. AS IT IS EITHER BACKGROUND
C         OR TAIL OF PRECEDING PEAK
      DO 300 I = 4,9
      IF ( Y(I) .LT. Y(I+1) ) GO TO 400
300   CONTINUE
      GHGT = 0
      RETURN
400   CONTINUE
      MI=M+39
C ARITHMETIC BELOW HAS BEEN REPLACED BY FUNCTION SUBROUTINE
C IN ORDER TO ACCOMODATE MAGNETIC INSTRUMENTS....
C                                         RTI WFH 11/20/78
C     TSHIFT = FLOAT(MI - LOMASS)/(HIMASS - LOMASS)
      TSHIFT = SCAN(MI)
C      INTERPOLATE PEAK PROFILE TO TIME COORDINATE OF MODEL PEAK.G
      CALL TIMESF(TSHIFT.Y(2).11.GMODE)
C      CANNOT USE ENDPOINTS AFTER INTERPOLATION
C
C
C      *** NOW USE LEAST SQUARES DOUBLET RESOLVER TO GET
C      *** THE AMPLITUDE OF THE COMPONENT AT THE DETECTED SPECTRUM.
C
      INSPEC = 6+(NPOS-1)/3
```

```
C        IHSPEC = 6 - 11
        STORE = 0
C        STORE WILL BE SET 1 BY DUBLET IF M QUALIFIES FOR NEW AMPLITUDE.
        CALL DUBLET(GNU.HNU.Y.STORE.M.IHSPEC.GHGT.ISW.NS)
C
C        SAVE THE COMPONENT CONTRIBUTING TO  THE DETECTED SPECTRUM.
C
C
C        SAVE CONTRIBUTION WITH MINIMUM ERROR.
C
        IF (GHGT .GT. 32767) GHGT = 32767
C        MADE SURE NO INTEGER OVERFLOW
        IF (GHGT .LE. 10) GHGT = 0
C        HEIGHTS LE 10 ARE CONSIDERED ARTIFACTS OF THE DOUBLET RESOLVER
        IF(STORE .EQ. 1) NUSPEC(M)=GHGT
C
5       RETURN
        END
C        THIS SUBROUTINE DOES ALL PRINTING
C        FOR CLEANUP. PLEASE USE IT TO ADD NEW PRINTING
C        IN ANY SUBROUTINE.
C        THE CALL STATEMENT ALLOWS FOR PASSING OF ARRAYS
C        REALS AND INTEGERS.
C        THIS COMPUTER PROGRAM WAS DEVELOPED WITH THE FUNDING SUPPORT
C        FROM THE NATIONAL INSTITUTES OF HEALTH (GRANTS RR-612 AND
C        GM-20832) AND THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
C        (GRANT NGR-05-020-632).
        SUBROUTINE SPEPRT(N.IAR.AR.R1.R2.R3)
        INTEGER N.IONE.NINE.MAD
        REAL R1.R2.R3
C THE DIMMENSIONS OF AR AND IAR HAVE BEEN SET TO (1).(1) FROM
C TO HELP WITH WATFIV....
C
        DIMENSION IAR(1).AR(1)                        WFH RTI 4/30/78
        COMMON /DUBBUG/IFLAG
        COMMON /IMAD/MAD
        COMMON /UNITS/ IRDR.IPTR1.IPTR2.IPTR3.IIN.IOUT
        DATA IONE.NINE/1.9/
C        IFLAG IS RESET IF NO PRINTING IS TO BE DONE/
        IF (IFLAG.EQ.0) RETURN
        GO TO (40.50.60.70).N
C        THE ABOVE IS A CASE STATEMENT FOR EXECUTING PRINT
C        OPTIONS.C
C        NOW DO STALSD OUTPUT
40      IIAR=IAR(1)-MAD
        WRITE(IPTR2.600) IIAR
        RETURN
C        THE ABOVE WRITE IS FOR MASS OF SATURATION
50      WRITE(IPTR2.550) R1
        WRITE(IPTR2.610)
        IUP = IAR(1)
        WRITE(IPTR2.535) (AR(I).I=IONE.IUP)
        RETURN
60      WRITE(IPTR2.800)
        WRITE(IPTR2.560) AR(1).R1.R2.R3
        WRITE(IPTR2.900)
C        WRITES RPKHGT.LOCAL BACKGROUND.RATIO.G(5)
        RETURN
```

```
70    WRITE(IPTR2.700)
C      THE ABOVE WRITES SATURATED DEFAULT VALUE.
      RETURN
C      FORMAT STATEMENTS FOLLOW.
535   FORMAT(/.8X.7F7.0./.8X.6F7.0./)
550   FORMAT(8X./'    SATURATED VALUE = '.F7.1)
560   FORMAT(8X.' PEAK HEIGHT ='.F7.0./.8X.' BACK GROUND ='.F7.0.
     -/.8X.' PEAK TO MODLEL RATIOS AVERAGE. RATIO ='.E15.8./.8X.
     -/.' G(5) ='.F7.0.//)
600   FORMAT(/.' SATURATION ENCOUNTERED. MASS = '.I4./)
610   FORMAT(' SATURATED PEAK PROFILE FOLLOWS  '/)
700   FORMAT(/.' SUPER SATURATION ENCOUNTERED './.8X.
     -'PEAK HEIGHT DEFAULTS TO MAXIMUM VALUE '//)
800   FORMAT(/' SATURATED PEAK ANAYLSIS   '/)
  900 FORMAT(/' NOTE: PEAK HEIGHT = RATIO X G(5)'/)
      END
      SUBROUTINE TIMESF(DELTIM.Y.N.GTIME)
      REAL    GTIME
      INTEGER N
      REAL    RATE.DELTIM
C Y HAS BEEN REDIMENSIONED DOWN FROM 15 TO 1
C TO MAKE THIS THING WATFIV COMPATABLE  -  WFH RTI 4/25/78
      DIMENSION    Y(1)
C
C       THIS COMPUTER PROGRAM WAS DEVELOPED WITH THE FUNDING SUPPORT
C       FROM THE NATIONAL INSTITUTES OF HEALTH (GRANTS RR-612 AND
C       GM-20832) AND THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
C       (GRANT NGR-05-020-632).
C
C       THIS SUBROUTINE TIME ALIGNS A PEAK AT A GIVEN MASS  M  TO
C       THE SAME TIME CO-ORDINATE SYSTEM AS THE MODEL'S  MASS.
C       (THE TIME CO-ORDINATE  IS  GTIME FOR MODEL)
C
C              DELTIM=(M-LOMASS)/(HIMASS-LOMASS)
C
C              Y IS THE ABCISSAS OF THE FRAGMENTOGRAM AT MASS  M.
C
C       THIS ALIGNMENT IS DONE BY A CUBIC FIT USING  LAGRANGE'S
C       METHOD.I.E.. GIVEN THE 4 POINTS  X0.X1.X2.X3. WE FIND L(X) AS
C       FOLLOWS:
C
C              L(X) =*(X-X1)(X-X2)(X-X3)/((X0-X1)(X0-X2)(X0-X3))@Y0 +
C                    *(X-X0)(X-X2)(X-X3)/((X1-X0)(X1-X2)(X1-X3))@Y1 +
C                    *(X-X0)(X-X1)(X-X3)/((X2-X0)(X2-X1)(X2-X3))@Y2 +
C                    *(X-X0)(X-X1)(X-X2)/((X3-X0)(X3-X1)(X3-X2))@Y3
C
C              WITH THE X'S INTEGRALLY SPACED.
C
C       IN EACH CASE WE INTERPOLATE FOR A NEW MIDDLE POINT.OR
C       L(X1+DX). SO. THE ABOVE EQUATION REDUCES TO
C
C              L(X1+DX) = C0 * Y0 + C1 * Y1 + C2 * Y2 + C3 * Y3
C
C          C0 = DX * (DX-1)(DX-2)/-6
C       C1 = (DX+1)(DX-1)(DX-2)/2
C       C2 = (DX+1)(DX)(DX-2)/-2
C       C3 = (DX+1)(DX)(DX-1)/6
C
```

```
C      WHERE DX = GTIME  -  DELTIM
C
      INTEGER I
      REAL    C1.C2.C3.DTIME
      REAL    YI.YI1.YMID.K1.K2.K3.K4.DXPOS
      DIMENSION YOLD(15)
C
C
C
      DO 50 I=1.N
      YOLD(I) = Y(I)
50    CONTINUE
C
C      SAVED OLDIES
C
      DX = GTIME - DELTIM
      DXPOS = DX
      IF ( DXPOS .LT. 0 ) DXPOS = -DX
      IF ( DXPOS .LT. 0.00001) RETURN
C      DON'T INTERPOLATE FOR SMALL VALUES OF DX
      K1 = DX - 1
      K2 = DX - 2
      K3 = DX + 1
      K4 = DX + 2
      C1= (DX * (K1) * (K2))/(-6)
      C2= ((K3) * (K1) * (K2))/2
      C3= ((K3) * (DX) * (K2))/(-2)
      C4= ((K3) * (DX) * (K1))/6
C
C      INTERPOLATE
C
      JN = N - 2
      DO 100 I=2.JN
      Y(I)=C1*YOLD(I-1) + C2*YOLD(I) + C3*YOLD(I+1) + C4 * YOLD(I+2)
100   CONTINUE
      E1= C4
      E2= ((K4) * (DX) * (K1))/2
      E3= ((K4) * (K3) * (K1))/(-2)
      E4= ((K4) * (K3) * (DX))/6
      NN = N - 1
      Y(NN) = Y(N-3) * E1 + Y(N-2) * E2 + Y(NN) * E3 + Y(N) * E4
C
      DO 275 I = 2.NN
      IF(Y(I) .LT. 0) Y(I) = 0
275   CONTINUE
      RETURN
      END
      SUBROUTINE TLOC(NS.M1.M2.M3.IFOUND.EOF)
C      THIS COMPUTER PROGRAM WAS DEVELOPED WITH THE FUNDING SUPPORT
C      FROM THE NATIONAL INSTITUTES OF HEALTH (GRANTS RR-612 AND
C      GM-20832) AND THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
C      (GRANT NGR-05-020-632).
C      *** THE MAIN FUNCTION OF THIS ROUTINE IS TO EXAMINE
C      *** THE CURRENT DATA IN THE PROCESSING WINDOW
C      *** AND DETECT WHEN THERE IS A COMPONENT PRESENT IN THE
C      *** WINDOW. THIS IS DONE BY KEEPING A HISTOGRAM OF
C      *** HOW MANY PEAK MODES ARE MAXIMIZING WITHIN
C      *** THE WINDOW. A HISTOGRAM IS ALSO KEPT OF THE
```

```
C      *** ION CURRENTS ASSOCIATED WITH THE PEAKS THAT ARE
C      *** ARE MAXIMIZING. THIS LATTER MEASURE IS MORE
C      *** EFFECTIVE FOR COMPONENT DETECTION. AN INDEX
C      *** SET IS KEPT OF THOSE MASSES THAT HAVE MAXIMIZED
C      *** IN THE WINDOW (ARRAY 'IND'). DURING THIS WHOLE
C      *** PROCESS A MODEL PEAK FOR THE COMPONENT IS ALSO
C      *** CHOSEN 'GPEAK' USING A SHARPNESS CRITERION.
C PARAMETER EOF ADDED - SEE COMMENT NEAR CALL STMT IN MAIN
C                                                 WFH RTI 9/31/78
      INTEGER NAME.RDATE.BITMAP.GCTIME.LIBMAP.INTEG.LOMASS
      INTEGER HIMASS.POINTS.TAPCNT.MAXSP.LSTCNT.DEVICE.SPCNO
      INTEGER SPEC.IND.IR.BK.NM.NTM.RECNO.FLAG.KSTOR(50).MAD.
     1KMAD(2.50)
      INTEGER NUSPEC.NMASS.OVLAP.EOF
      INTEGER MODBKG.LOCLBG
      INTEGER GPEAK.SPARE1.SPARE2.SPARE3.SPARE4.SPARE7
      INTEGER  FSTREC.AJACEN.AHEAD
      REAL    SPARE5.SPARE6
C THIS HAS BEEN INSERTED FOR THE AVERAGING FACILITY
C                                          RTI WFH 1/14/79
      INTEGER NAVG.HAVG.AVGSPC(411)
      REAL*8 TLC/' TLOCXX '/
      REAL     ISVP8.ISVP9
C THE IDENTIFIER DIB HAS BEEN CHANGED TO SCAN TO MAKE
C WAY FOR THE NEWLY INCORPORATED FUNCTION SUBROUTINE
C                                       WFH RTI 9/31/78
      REAL    F1.F2.F3.YTOP.HIST(7).SLOPE.SCAN
      COMMON PKHIST(20).NPHIST(20)
      COMMON YM5(11)
      COMMON GM4.GM3.GM2.GM1.G0.GP1.GP2.GP3.GP4
      COMMON SG.SG2.SIG.SI.SI2.AN.AA.BB.CC.DD
      COMMON NAME(32).RDATE(5).BITMAP(64).GCTIME
      COMMON LIBMAP(17).INTEG.LOMASS.HIMASS.POINTS.TAPCNT
      COMMON MAXSP.LSTCNT.DEVICE(5).SPCNO
      COMMON SPEC(14.411).IND(3.411).NUSPEC(411).NMASS
      COMMON RMAX(3).IR(3).BK(3).NM(3).NTM(3).GPEAK(3.11)
      COMMON SPARE1.SPARE2.SPARE3.SPARE4.SPARE5.SPARE6.SPARE7
      COMMON RATMIN.RATEMX.SIGY0.SIGNXT.SIGLST
      COMMON ISVN8.ISVN9.ISVP8.ISVP9.SIGMA7.SATVAL.SATMAX
      COMMON MODE1.MODE2.NLWCUT.NUPCUT.SATPKS(25).NSAT.EXPNAM(5)
      COMMON /DUBBUG/IDEBUG
      COMMON /MODES/SGMODE(3).SHMODE(3).GMODE.HMODE
      COMMON /MASSES/ISGMAS(3).ISHMAS(3).IGMASS.IHMASS
      COMMON /IFF/IFLAG(3)
      COMMON /AJACNT/AJACEN(2)
      COMMON /PARAMS/MINTIC.MINNY.NTHIRD.RATM
      COMMON /LFTRIT/ILEFT.IRIGHT
      COMMON /HIND/IH1.IH2.MSINGS(2.50).MINX.LINGS(50).LINX
C     COMMON /CRAWBU/RAWBUF(411)
      COMMON /CFSTRE/FSTREC
      COMMON /AVG/ NAVG.HAVG
      COMMON /CAHEA/AHEAD(5)
      COMMON /IMAD/MAD
C      MSINGS SAVES MASSES  OF SINGLETS IN  HISTOGRAMS.
      EQUIVALENCE (YM4.YM5(2))
      EQUIVALENCE (YM3.YM5(3)).(YM2.YM5(4)).(YM1.YM5(5))
      EQUIVALENCE (Y0.YM5(6)).(YP1.YM5(7)).(YP2.YM5(8))
      EQUIVALENCE (YP3.YM5(9)).(YP4.YM5(10)).(YP5.YM5(11))
      DATA HIST/-.3333,0,.3333,.6667,1,0,1.3333,1.6667/
C      *** CIRCULAR QUEUE TO CONTROL CURRENT DATA SET
      I1=NS-13-14*((NS-14)/14)
      I2=NS-12-14*((NS-13)/14)
      I3=NS-11-14*((NS-12)/14)
      I4=NS-10-14*((NS-11)/14)
      I5=NS-9-14*((NS-10)/14)
      I6=NS-8-14*((NS-9)/14)
      I7=NS-7-14*((NS-8)/14)
      I8=NS-6-14*((NS-7)/14)
      I9=NS-5-14*((NS-6)/14)
      I10=NS-4-14*((NS-5)/14)
      I11=NS-3-14*((NS-4)/14)
      I12=NS-2-14*((NS-3)/14)
      I13=NS-1-14*((NS-2)/14)
      I14=NS-14*((NS-1)/14)
C      *** READ IN NEXT SPECTRUM
C      NOTE: NS = 14,15,....,MAXSPEC - 10
C            FIRST RECORD TO BE READ IS NUMBER 13.
      RECNO = NS
      INSLT=1
C
C THE PRIMARY INPUT STATEMENT HAS BEEN MODIFIED TO INCLUDE
C AN AVERAGING FACILITY OVER NAVG INPUT SPECTRA
C                                           RTI - WFH 1/13/79
C INITIALIZE CURRENT SPECTRUM IN CORE
      DO 90 IK=1,411
   90 SPEC(I14,IK)=0
C ADD AS MANY INPUT SPECTRA AS NECESSARY TO CORE SPECTRUM
      DO 92 IJ=1,NAVG
C IQUPD ENSURES THAT THE QUEUE DOWN INSIDE THE READ ROUTINES IS
C UPDATED ONLY ONCE OVER THE SPREAD OF NAVG INPUT SPECTRA
      IQUPD=0
      IF(IJ.EQ.HAVG)IQUPD=1
C IQUPD TELLS WETHER OR NOT TO UPDATE READ ROUTINE QUEUE
C I14 TELLS RDSPC QUEUE POSITION
      CALL RDSPC(ISCNO,AVGSPC,I14,IQUPD,EOF)
      IQUPD=0
      IF(EOF.EQ.0)GO TO 4
      RETURN
C READ ONWARD...
    4 CONTINUE
      DO 97 IK=1,411
   97 SPEC(I14,IK)=SPEC(I14,IK)+AVGSPC(IK)
   92 CONTINUE
C DIVIDE AS NECESSARY
      IF(NAVG.EQ.1)GO TO 94
      DO 95 IK=1,411
   95 SPEC(I14,IK)=SPEC(I14,IK)/NAVG
   94 CONTINUE
C NOW UPDATE SPECTRUM NUMBER
      SPCNO=ISCNO-7*NAVG+HAVG-1
C THE FOLLOWING TEN STMENTS OR SO HAVE BEEN DISCARDED
C TO MAKE WAY FOR MULTIPLE RUN PROCESSING AS WELL
C AS THE COMMON BLOCK /CRAWBU/ AS NEITHER SERVES ANY
C USEFULL PURPOSE.
C
C                               WFH - RTI
C                               12/6/78
```

```
C     GO TO 620
C2    WRITE(IPTR2.31)
C   2 CALL TLPRNT(3.NPHIST.PKHIST.XY.XY.XY.1.1.1)
C     STOP 16
C  31 FORMAT(' ***FATAL READ ERROR ON MAIN INPUT FILE-IN TLOC***')
C     RETURN
C620  DO 10 M=1.NMASS
C10   SPEC(114.M) = RAWBUF(M)
C       NOW.  WE'RE OFF AND RUNNING WHILE THE DISK IS HUMMING ....
      NS3=NS-11
      M3=NS3-3*((NS3-1)/3)
      M2=NS3-1-3*((NS3-2)/3)
      M1=NS3-2-3*((NS3-3)/3)
      RMAX(M3)=0.0
      SGMODE(M3)=0
      BK(M3)=0
      IR(M3)=0
      ISGMAS(M3) = -1
      NMAX=0
      SIGMA7=0.0
      IFLAG(M3) =  0
      NAVE = 0
C      LINGS IS SET UP IN NXTMAX AND CONTAINS STUFF FROM HIST 10.11
C      WHICH BECOMES HIST OF 7.8 HERE.
      DO 12 I = 1.50
      MSINGS(IH1.I) = MSINGS(IH2.I)
12    MSINGS(IH2.I) = LINGS(I)
C      SET MINX FOR STORING IN MSINGS.
      DO 13 I = 1.49
      IF (LINGS(I) .EQ. 0) GO TO 14
13    CONTINUE
14    MINX = I
      LINX = 1
      DO 17 J = 1.1
17    LINGS(J) = 0
      NINX = 1
      DO 18 I = 1.50
      IF(MSINGS(IH1.I) .EQ. 0 ) GO TO 19
18    CONTINUE
19    CONTINUE
      NINX = I
C      OINX IS USED FOR STORING HISTOGRAMED MASSES IN H(5-6)
C      NOW MOVE DETECTION FLAG FOR PREVIOUS SPECTRUM INTO
C      THE FLAG FOR ITS IMMEDIATE PREDECESSOR.THIS IS USED
C      IN GETHGT SO SINGLETS FROM IMMEDIATED PREDECESSORS
C      ARE NOT INCLUDED IN THEIR IMMEDIATE SUCESSORS.
      AJACEN(1) = AJACEN(2)
      AJACEN(2) = 0
C       LOOP THROUGH ALL MASSES       AND RESET FOUND INDICES.
      DO 15 M=1.411
15    IND(M3.M)=0
      NPHIST(8)= ISVN8
      NPHIST(9)= ISVN9
      PKHIST(8)= ISVP8
      PKHIST(9)= ISVP9
      DO 16 J=1.6
      PKHIST(J)=PKHIST(J+3)
16    CONTINUE
```

```
C        PKHIST(1)←PKHIST(4).....PSHIST(6)←PKHIST(9)
       PKHIST(7)= (SPARE5)
       PKHIST(8)= (SPARE6)
       PKHIST(9)=0
       DO 231 J=1,6
       NPHIST(J)=NPHIST(J+3)
231    CONTINUE
C        NPHIST(1)←NPHIST(4).....NPHIST(6)←NPHIST(9)
       NPHIST(7)= SPARE3
       NPHIST(8)= SPARE4
       NPHIST(9)=0
       NPHIST(10)=0
       NPHIST(11)=0
       PKHIST(10)=0
       PKHIST(11)=0
       DO 43 I = 2,5
43     AHEAD(I-1) = AHEAD(I)
       AHEAD(5) = 0
C       THE ABOVE IS A BINARY LIST OF SPECTRA DETECTED AHEAD.
C       WE SHIFT IT LEFT AT EACH NEW SPECTRUM SEARCH.IT IS
C       USED TO GUARANTEE ANY SPECTRUM DETECTED AHEAD IS ALWAYS
C       PROCESSED.
C STATEMENT REMOVED BECAUSE NO LONGER NEEDED WITH SCAN ROUTINE
C                                            WFH RTI 9/31/78
C      DIB = HIMASS - LOMASS
       DO 70 M=1,411
C       RUN THROUGH MASSES
       Y0=SPEC(I8,M)
C       SAVE FOR CLEAN UP RATIO. TICRAT
       SIGMA7 = SIGMA7 + Y0
       IF(Y0.GE.SATVAL) GO TO 65
C       *** MARK SATURATED MASSES FOR LATER PROCESSING
       YM1=SPEC(I7,M)
       YP1=SPEC(I9,M)
C       TEST FOR SINGLET
C           .  .
C          .      .
C       .    Y0         .              <-- SHAPE REQUIREMENT.
C      .                     .
C  .
C                    YP4
       IF(Y0 .LE. YM1) GO TO 70
       IF(Y0 . LE. YP1) GO TO 70
       YM2=SPEC(I6,M)
       YP2=SPEC(I10,M)
       YM3=SPEC(I5,M)
       YP3=SPEC(I11,M)
       YM4 =SPEC(I4,M)
       YP4=SPEC(I12,M)
       YM5(1)=SPEC(I3,M)
       YP5=SPEC(I13,M)
       YY5=YM5(1)
C       TEST LEFT SLOPE OF PEAK
C       ILEFT IS REQUIRED NUMBER OF STRICKLY DECREASING LEFT POINTS
       IF(YM1 .LE. YM2) GO TO 65
       IF(ILEFT .EQ. 2) GO TO 44
       IF(YM2 .LE. YM3) GO TO 65
       IF(ILEFT .EQ. 3) GO TO 44
```

```
      IF(YM3 .LE. YM4) GO TO 65
44    CONTINUE
C      TEST RIGHT SLOPE OF PEAK
C      IRIGHT IS REQUIRED NUMBER OF STRICKLY DECREASING RIGHT POINTS
      IF(YP1 .LE. YP2) GO TO 65
      IF(IRIGHT .EQ. 2) GO TO 45
      IF(YP2 .LE. YP3) GO TO 65
      IF(IRIGHT .EQ. 3) GO TO 45
      IF(YP3 .LE. YP4) GO TO 65
45    CONTINUE
C       CALCULATE VALUE AT MODE OF PEAK,YTOP, AND
C        X DISPLACEMENT,XTOP.
      CALL NEWTOP(YM5,RATE,XTOP,YTOP,4)
C      IF XTOP IS OUT OF RANGE THE PEAK IS NOT ELUTING HERE OR
C      THE DATA HAS TOO MANY INFLECTION POINTS
      IF(XTOP .LE. -1.  .OR.  XTOP .GE. 1.) GO TO 65
C      MAKE SURE THE PEAK IS INCREASING FASTER THAN BACKGROUND
      IF( (YTOP - YY5).LT.30  .AND.  (YTOP - YP5).LT.30 ) GOTO 70
C      SEE IF QUALIFIES FOR MODEL FROM RATE.
      IF(YTOP.GE.SATVAL.OR.YP1.GE.SATVAL.OR.YM1.GE.SATVAL)GO TO 65
      IF(YP2.GE.SATVAL.OR.YP3.GE.SATVAL.OR.YP4.GE.SATVAL) GO TO 65
      IF(YM2.GE.SATVAL.OR.YM3.GE.SATVAL.OR.YM4.GE.SATVAL)GO TO 65
C      PREVENT SUPER SATURATION BY ABOVE
C      SEE IF RATE IS MAXIMUM TO QUALIFY AS MODEL PEAK
      MODEL = 0
      MI = M + 39
      SLOPE = (YP5 - YY5)/10
      LOCLBG = YY5 + SLOPE * 5
C CHANGED TO HAVE FUNCTION SUBROUTINE TO HAVE DIFFERENT INSTRUMENTS
C                                         WFH RTI 9/31/78
      TSHIFT=SCAN(MI)
C     TSHIFT = (MI - LOMASS)/DIB
      IF(RATE.LT.RMAX(M3)) GO TO 50
      MODEL = 1
      IF((YTOP - YY5) .GE. 160) GO TO 2000
      IFLAG(M3) = 1
C      FLAG AS FUNNY MODEL
      GO TO 2500
2000  IFLAG(M3) = 0
2500  CONTINUE
      RMAX(M3)=RATE
      IR(M3)=M
C      USE LINEAR BACKGROUND AS FIRST CUT APPROXIMATION
      MODBKG = SPEC(I1,M)
      IF ( MODBKG .GT. SPEC(I2,M) ) MODBKG = SPEC(I2,M)
      IF ( MODBKG .GT. SPEC(I3,M) ) MODBKG = SPEC(I3,M)
      IF ( MODBKG .GT. SPEC(I4,M) ) MODBKG = SPEC(I4,M)
      IF ( MODBKG .GT. SPEC(I5,M) ) MODBKG = SPEC(I5,M)
      IF ( MODBKG .GT. SPEC(I11,M) ) MODBKG = SPEC(I11,M)
      IF ( MODBKG .GT. SPEC(I12,M) ) MODBKG = SPEC(I12,M)
      IF ( MODBKG .GT. SPEC(I13,M) ) MODBKG = SPEC(I13,M)
      IF ( MODBKG .GT. SPEC(I14,M) ) MODBKG = SPEC(I14,M)
C      MODBKG = MIN ( SPEC(IJ,M) : I1 .LE.  IJ  .LE. I14)
      DO  40   IJ = 1,11
      GPEAK(M3,IJ) = YM5(IJ) - MODBKG
40    CONTINUE
      ISGMAS(M3) = MI
      SGMODE(M3)=TSHIFT
      BK(M3) = MODBKG
50    CONTINUE
      XY=YTOP - (LOCLBG)
      IF ( XY .LE. 0 ) XY = 1.0
C SIGY0 (EXCEPT IN COMMON DECLARATIONS) APPEARS ONLY IN THIS
C STATEMENT. HENCE IT HAS BEEN COMMENTED OUT.
C                                      WFH AT RTI        4/19/78
C     SIGY0=SIGY0 +  XY
      TITTY = TSHIFT + XTOP
      DO 800 IJ = 1.7
      IF (TITTY .LT. HIST(IJ)) GO TO 900
800   CONTINUE
900   IH = 4 + IJ
C      IH = 5 TO 11
C      HISTOGRAM DATA PKHIST FOR INTENSITY. NPHIST FOR COUNT OF PEAKS
      PKHIST(IH)=PKHIST(IH) + XY
      NPHIST(IH)=NPHIST(IH)+1
C      SAVE MODE OF MODEL PEAK
C      RESET MODEL FLAG
      IF ( MODEL .EQ. 1 ) GO TO 925
C      ALWAYS LOG MASS OF MODEL.
      NGOGO = IH - 4
C      NGOGO = 1,2,3,4,5,6,7
      GO TO ( 950.950.925,925.925.960.960 ), NGOGO
C  SLOT                   5   6   7   8   9  10  11
C      HISTOGRAM SLOTS + 7.8.9 0
925   MSINGS(IH2,MINX) = MI
      MINX = MINX + 1
      IF(MINX .EQ. 51) MINX = 50
      IF ( MODEL .EQ. 1 ) GO TO 950
      GO TO 990
950   CONTINUE
C      HISTOGRAM SLOTS 5,6
      IF ( IH .GE. 7 ) GO TO 960
      MSINGS(IH1,NINX) = MI
      NINX = NINX + 1
      IF(NINX .EQ. 51) NINX = 50
      IF ( MODEL .EQ. 0 )  GO TO 990
C          WE HAVE A POSSIBLE MODEL FOR PREVIOUS SPECTRUM SO
C           IF THE RATE CRITERIA IS SATISFIED, SAVE IT.
      IF ( RATE .LE. RMAX(M2) ) GO TO 990
      ISGMAS(M2) = MI
      RMAX(M2) = RATE
      IFLAG(M2) = 0
      DO 970 IM = 1,11
970   GPEAK(M2,IM) = GPEAK(M3,IM)
C      HISTOGRAM SLOTS 10,11
960   CONTINUE
      IF ( IH .LT. 10 ) GO TO 990
      LINGS(LINX) = MI
      LINX = LINX + 1
      IF(LINX .EQ. 51) LINX = 50
990   CONTINUE
C      SAVE MAXIMUM OF 50 MASSES DUE TO CORE LIMITATIONS
65    CONTINUE
      NMAX=NMAX+1
C      SAVE MASS THAT IS CONSIDERED A CONTRIBUTION
C      NMAX IS NUMBER OF MASSES MAXIMIZING .
```

```
      IND(M3.M) = M
70    CONTINUE
C      NOW CHECK FOR DETECTED SPECTRUM IN *4.5.60
      NL=1
      NU=9
      IDXL=4
      IDXU=6
      SPARE3=NPHIST(10)
      SPARE4=NPHIST(11)
      SPARE5=PKHIST(10)
      SPARE6=PKHIST(11)
      ISVN8=NPHIST(8)
      ISVN9=NPHIST(9)
      ISVP8=PKHIST(8)
      ISVP9=PKHIST(9)
      CALL DETEKT(PKHIST.NPHIST.IFOUND.NLWCUT.NUPCUT
     &   .NL.NU.IDXL.IDXU.MODE1)
C      MODE1 IS THE LOCATION HISTOGRAMICALLY OF THE SPECTRUM.
C      4.LTE.MODE1.LTE.6
C      IFOUND IS 1 IF SPECTRUM DETECTED.
C PUT IN HEURISTIC FIX FOR THE CASE WHERE NO SUITABLE
C MODEL IS FOUND - INSERTED AT RTI BY WFH - 5/23/78
C
      IF(ISGMAS(M3).EQ.-1.AND.ISGMAS(M2).EQ.-1)IFOUND=0
      IF (IFOUND.NE.1) GO TO 888
      AJACEN(2) = 1
C      SETS SPECTRUM DETECTED FLAG.
C TEMPORARY MOD HAS BEEN ADDED TO GATHER STATISTICS ON
C PEAK RATES-                  RTI WFH 9/31/78
C     IOBS=1
C     WRITE(20.6251) IOBS.SPCNO.RMAX(M3)
C6251 FORMAT(A4.A4.12X.A4)
      IF (RMAX(M3) .GE. RATM) GO TO 889
      AJACEN(2) = 0
      IFOUND = 0
C..          CALL PUTNUM('SPEC. REJECTED FOR MIN. RATE. SPEC ='.SPCNO)
C..       CALL PUTFLT('    RATE = '.RMAX(M3))
      CALL TLPRNT(4.NPHIST.PKHIST.RMAX(M3).XY.XY.SPCNO.1.1)
C     WRITE(IPTR2.32)SPCNO.RMAX(M3)
C32   FORMAT(' SPEC. REJECTED FOR MIN. RATE. SPEC = '.I6/'  RATE = '.F10
C    *.3/)
      GO TO 888
889   CONTINUE
C      IFOUND = 1 IF SPECTRUM DETECTED.
      CALL TLPRNT(1.SPEC.SPEC.XY.XY.XY.SPCNO.NLWCUT.NUPCUT)
      CALL TLPRNT(2.NPHIST.PKHIST.XY.XY.XY.1.1.1)
      IF ( IDEBUG .EQ. 0 ) GO TO 888
      CALL TLPRNT(5.NPHIST.PKHIST.XY.XY.XY.1.1.1)
C     WRITE(IPTR2.5000)
      DO 9 KX=1.2
      DO 9 KZ=1.50
      KMAD(KX.KZ)=MSINGS(KX.KZ)-MAD
      IF(KMAD(KX.KZ).LE.0)KMAD(KX.KZ)=0
9     CONTINUE
      DO 885 I = 1.5
      JLOW = (I - 1) * 10 + 1
C      JLOW = 1.11.....41
      JUP = JLOW + 9
```

```
C       JUP = 10.20.....50
      IF (KMAD(IH1.JLOW) .LE. 0) GO TO 886
C YET ANOTHER MOD TO TLPRNT
C                                WFH RTI 10/1/78
      DO 731 IK=1.50
  731 KSTOR(IK)=KMAD(IH1.IK)
      CALL TLPRNT(6.KSTOR.PKHIST.XY.XY.XY.JLOW.JUP.1)
C     WRITE(IPTR2.5100) (KMAD(IH1.K). K = JLOW.JUP)
885   CONTINUE
886   CONTINUE
888   CONTINUE
      NM(M3)=NMAX
C       SAVE NUMBER OF MASSES MAXIMIZING
      RETURN
C5000 FORMAT(//.' SINGLET MASSES CONTRIBUTING TO DETECTED SPEC. (MAXIMUM
C    * OF 50 PRINTED) '/)
C5100 FORMAT(10(2X.I3))
      END
C    THIS ROUTINE DOES ALL PRINTING FOR CLEANUP.  PLEASE USE IT TO ADD
C    NEW PRINTING IN ANY SUBROUTINE.
C       THE CALL STATEMENT ALLOWS FOR PASSING OF ARRAYS
C       REALS AND INTEGERS.
C       THIS COMPUTER PROGRAM WAS DEVELOPED WITH THE FUNDING SUPPORT
C       FROM THE NATIONAL INSTITUTES OF HEALTH (GRANTS RR-612 AND
C       GM-20832) AND THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
C       (GRANT NGR-05-020-632).
      SUBROUTINE TLPRNT(N.IAR.AR.R1.R2.R3.I1.I2.I3)
      INTEGER N.I1.I2.I3.IONE.NINE.IP/0/
      REAL R1.R2.R3
      DIMENSION IAR(1).AR(1)
      COMMON /DUBBUG/IFLAG
      COMMON /UNITS/ IRDR.IPTR1.IPTR2.IPTR3.IIN.IOUT
      DATA IONE.NINE/1.9/
C       IFLAG IS RESET IF NO PRINTING IS TO BE DONE/
C     IF (IFLAG.EQ.0) RETURN
      GO TO (20.30.40.50.60.70.80) .N
C       THE ABOVE IS A CASE STATEMENT FOR EXECUTING PRINT
C20      WRITE(IPTR2.515)
   20 IF(IFLAG.EQ.0)RETURN
      WRITE(IPTR2.520) I1.I2.I3
C....WRITE TO TERMINAL
      IF(IP.EQ.1)RETURN
C     IF(IP.EQ.1)GO TO 21
C THE WRITE TO UNIT 7 STATEMENTS WERE DETERMINED TO BE
C SUPERFLUOUS AND WERE COMMENTED OUT 3/30/78 -
C                                            WFH
C     WRITE(7.521)
      IP=1
C 21  WRITE(7.22)I1
C 22  FORMAT(I6)
C      THIS DOES TLOC STUFF
      RETURN
   30 IF(IFLAG.EQ.0)RETURN
      WRITE(IPTR2.1400)
      WRITE(IPTR2.530) (IAR(I).I=IONE.NINE)
C       ABOVE WRITES NPHIST
      WRITE(IPTR2.535) (AR(I).I=IONE.NINE)
C       ABOVE  WRITES  PKHIST
```

```
      RETURN
C515     FORMAT(/.10('-------'))
520   FORMAT(1H1.1X.'  ****** SPECTRUM DETECTED ******'./.
     -/' SPECTRUM NO. IS '.I5.'.   LEFT AND RIGHT HISTOGRAM BOUNDS '.
     -'ARE '.I5.' AND'.I5.'.'/)
530   FORMAT(/.' PEAK HIST  '/.8X.9I6.//' ION CURRENT HIST ')
535   FORMAT(1X.9F9.0)
1400  FORMAT(' HISTOGRAM POSITIONS 1 - 9 ...'/)
C THESE MODS HAVE BEEN ADDED WITH ADDITIONAL STATEMENTS (WRITE)
C BEING INCORPORATED IN TLOC
C                                            WFH RTI 9/31/78
   40 WRITE(IPTR2.31)
   31 FORMAT(' ***FATAL READ ERROR ON MAIN INPUT FILE IN TLOC***')
      RETURN
   50 IF(IFLAG.EQ.0)RETURN
      WRITE(IPTR2.32) I1.R1
 32   FORMAT(' SPEC. REJECTED FOR MIN. RATE. SPEC = '.I6/'  RATE = '.F10
     *.3/)
      RETURN
   60 IF(IFLAG.EQ.0)RETURN
      WRITE(IPTR2.5000)
5000  FORMAT(//.' SINGLET MASSES CONTRIBUTING TO DETECTED SPEC. (MAXIMUM
     * OF 50 PRINTED) '/)
      RETURN
C
   70 IF(IFLAG.EQ.0)RETURN
      WRITE(IPTR2.5100)(IAR(I).I=I1.I2)
 5100 FORMAT(10(2X.I3))
      RETURN
C
   80 WRITE(IPTR2.55)
   55 FORMAT(1X.'INCORRECT NUMBER OF RECORDS SPECIFIED - '.
     1'EXECUTION NOT AFFECTED')
      RETURN
      END
      SUBROUTINE XTREM(ARRAY.NPTS.XMIN.MININD.XMAX.MAXIND)
      INTEGER ARRAY(411).XMIN.MININD.XMAX.MAXIND
      XMIN=ARRAY(1)
      XMAX=ARRAY(1)
      MININD=1
      MAXIND=1
      IF(NPTS.LE.1)RETURN
      DO 2 I=2.NPTS
      IF(ARRAY(I).GE.XMIN)GO TO 1
      XMIN=ARRAY(I)
      MININD=I
      GO TO 2
1     IF(ARRAY(I).LE.XMAX)GO TO 2
      XMAX=ARRAY(I)
      MAXIND=I
2     CONTINUE
      RETURN
      END
      SUBROUTINE OPNFIL
      INTEGER CLNFLG
      INTEGER MASBUF(6).VALBUF(6).AJACEN.AHEAD.ISO(1000)
      INTEGER NAME.RDATE.BITMAP.GCTIME.LIBMAP.INTEG.LOMASS.MAD.ADS
      INTEGER HIMASS.POINTS.TAPCNT.MAXSP.LSTCNT.DEVICE.SPCNO
      INTEGER LSTMIN.SPEC.IND.IR.BK.NM.NTM.PNAME(32)
      INTEGER HIST(10).UNIT.NNAME(18).IRNO
      INTEGER NUSPEC.NMASS.OVLAP.EXPNAM.SRLIB(5).SECOND
      INTEGER ODUBLT.OSATS.NLINE.RAWBUF.FSTPGE.FSTREC
      INTEGER SATLST.SATINX.PASS1.ITOM.EOF
      INTEGER GPEAK.SPARE1.SPARE2.SPARE3.SPARE4.SPARE7
      INTEGER*4 TIMREM.TIMCNT
      REAL     SPARE5.SPARE6
      REAL     ISVP8.ISVP9
      REAL    QNORM
      REAL*8    C4.C1.C2.SGX.SHX
      DIMENSION G(9).H(11).GNU(18).HNU(18)
      COMMON PKHIST(20).NPHIST(20)
      COMMON YM5.YM4.YM3.YM2.YM1.Y0.YP1.YP2.YP3.YP4.YP5
      COMMON GM4.GM3.GM2.GM1.G0.GP1.GP2.GP3.GP4
      COMMON SG.SG2.SIG.SI.SI2.AN.AA.BB.CC.DD
      COMMON NAME(32).RDATE(5).BITMAP(64).GCTIME
      COMMON LIBMAP(17).INTEG.LOMASS.HIMASS.POINTS.TAPCNT
      COMMON MAXSP.LSTCNT.DEVICE(5).SPCNO
      COMMON SPEC(14.411).IND(3.411).NUSPEC(411).NMASS
      COMMON RMAX(3).IR(3).BK(3).NM(3).NTM(3).GPEAK(3.11)
      COMMON SPARE1.SPARE2.SPARE3.SPARE4.SPARE5.SPARE6.SPARE7
      COMMON RATMIN.RATEMX.SIGY0.SIGNXT.SIGLST
      COMMON ISVN8.ISVN9.ISVP8.ISVP9.SIGMA7.SATVAL.SATMAX
      COMMON MODE1.MODE2.NLWCUT.NUPCUT.SATPKS(25).NSAT.EXPNAM(5)
      COMMON /UNITS/ IRDR.IPTR1.IPTR2.IPTR3.IIN.IOUT
C THIS COMPRISES THE FIRST OF SIX SUBROUTINES FOR THE EPA I/O
C INTERFACE TO THE CLEANUP PROGRAM. IN THE RTI SYSTEM. THE
C PRIMARY INPUT AND OUTPUT FILES WOULD BE OPENED HERE. FOR EPA
C ONLY A DUMMY RETURN IS NECESSARY.
C                                          WM.F.HARGROVE
C                                          RTI 11/25/78
      EQUIVALENCE (RSEC.ISEC)
C NOW PICK UP SATURATION VALUE FROM TTY...
      WRITE(IPTR1.3110)
 3110 FORMAT(1X.'ENTER SATURATION VALUE IN F6.0:')
      READ(IRDR.3112) SATVAL
 3112 FORMAT(F6.0)
      WRITE(IPTR1.3113) SATVAL
 3113 FORMAT(1X.'SATVAL= '.F8.1)
      RETURN
      END
      SUBROUTINE CLSFIL
C THIS COMPRISES THE SECOND OF SIX SUBROUTINES FOR THE EPA I/O
C INTERFACE TO THE CLEANUP PROGRAM. IN THE SYSTEM. THE
C PRIMARY INPUT AND OUTPUT FILES WOULD BE CLOSED HERE. FOR EPA
C ONLY A DUMMY RETURN IS NECESSARY.
C                                                     WM.F.HARGROVE
C                                                     RTI 11/25/78
      RETURN
      END
      SUBROUTINE CLSRUN(ISQ.MU)
C THIS COMPRISES THE THIRD OF SIX SUBROUTINES FOR THE EPA I/O
C INTERFACE TO THE CLEANUP PROGRAM. IN THE RTI SYSTEM. THE
C FIRST. LAST. AND NUMBER OF SCANS WOULD BE INSERTED INTO THE HEADER
C RECORD WHICH WOULD THEN BE WRITTEN OUT. FOR EPA ALL THAT IS NEEDED
C IS A RETURN.
      INTEGER MU.ISQ(1000)
      RETURN
      END
      SUBROUTINE OPNRUN(LIBBLK.IEOF)
C THIS COMPRISES THE FOURTH OF SIX SUBROUTINES FOR THE EPA I/O INTERFACE
C TO THE CLEANUP PROGRAM. THE HEADER RECORDS ARE READ IN
C AND WRITTEN OUT AGAIN AND THE CIRCULAR QUEUE FOR KEEPING TRACK OF SCAN
C TIMES IS INITIALIZED. THE MAIN DIFFERENCE BETWEEN EPA'S OPNRUN AND
C RTI'S OPNRUN IS THAT HERE FOR EPA SATVAL IS READ IN FROM THE TTY.
C WHEREAS RTI'S SETUP PICKS IT OFF THE RUN HEADER RECORD.
C                                                   WM.F.HARGROVE
C                                                   RTI 11/25/78
      INTEGER LIBBLK(1024).EOF.NS
      INTEGER IRDR.IPTR1.IPTR2.IPTR3.IIN.IOUT.IEOF.ISEC
      REAL SATVAL.SECSCN.RSEC
      INTEGER NAME(3).HDRFLG.DATE(2).RUNHDR.RUNMIN.
     1SAMPID(16).INAM(2).RUNCON(16).BUF4(11).LOPMAS.HIPMAS.
     2RETMIN(14).RETSEC(14).BLANK/'    './.INCOS(2)/'INCO'.'S   './
      COMMON /HDR1/ NAME.HDRFLG.DATE.RUNHDR.RUNMIN
      COMMON /HDR2/ SAMPID.INAM
      COMMON /HDR3/ RUNCON.SECSCN
      COMMON /HDR4/ BUF4.LOPMAS.HIPMAS
      COMMON /FLAGS/ EOF
C
      COMMON /UNITS/ IRDR.IPTR1.IPTR2.IPTR3.IIN.IOUT
      COMMON /CQUEUE/ RETMIN.RETSEC.NS
C INITIALIZE CIRCULAR QUEUE FOR SAVING SCAN TIMES
      NS=0
C HAND CONTROL TO PADDOCK
      CALL EPAHDR
      IEOF=EOF
      IF(IEOF.EQ.1)RETURN
C NOW PUT THINGS IN THEIR APPROPRIATE PLACES FOR CLEANUP
      IF(HDRFLG.EQ.0)GO TO 10
      WRITE(IPTR1.20)
   20 FORMAT(1X.'*** EXEXUTION TERMINATING. INCORRECT FORMAT OF HEADER
     1RECORD DETECTED AT BEGINNING OF FILE IN OPNFIL***')
      STOP
C NOW PICK UP DATE
   10 CONTINUE
      DO 30 I=1.2
   30 LIBBLK(I+30)=DATE(I)
C SET UP INSTRUMENT NAME
      DO 35 I=10.12
   35 LIBBLK(I)=BLANK
      DO 40 I=1.2
   40 LIBBLK(I+7)=INCOS(I)
C RUN SAMPLE ID
      LIBBLK(29)=BLANK
      LIBBLK(30)=BLANK
      DO 50 I=1.16
   50 LIBBLK(I+12)=SAMPID(I)
C REMARKS ON RUN CONDITIONS
      DO 60 I=1.16
   60 LIBBLK(I+137)=RUNCON(I)
      LIBBLK(155)=BLANK
      LIBBLK(154)=BLANK
C PICK UP SCAN RATE
      RSEC=SECSCN
```

```
C       LIBBLK(57)=ISEC
C MOVE IN MASS SCAN RANGE
        LIBBLK(42)=LOPMAS
        LIBBLK(43)=HIPMAS
C SET HDRFLG = FIRST SPECTRUM NUMBER IN RUN
        LIBBLK(36)=HDRFLG
C SET MAXIMUM NUMBER OF SPECTRA IN RUN
        LIBBLK(38)=100000
        RETURN
        END
        SUBROUTINE RDSPC(SPCNO.AVGSPC.I14.IQUPD.IEOF)
C THIS COMPRISES THE FIFTH OF SIX SUBROUTINES FOR EPA I/O
C INTERFACE TO THE CLEANUP PROGRAM. HERE IS WHERE CLEANUP
C REQUESTS THAT A SCAN RECORD BE READ. ASIDE FROM THE I/O
C STATEMENTS. THE LOGIC IS THE SAME AS FOR THE RTI LINEAR
C SYSTEM.
C
C                                               WM.F.HARGROVE
                                                11/25/78
      INTEGER AVGSPC(411).SPCNO.I14.EOF.IEOF
      INTEGER INAM(3).ISCN.IDBUF(2).IDHDR.IDMIN.IRTMIN.
     1IRTSEC.INMBAS.INBUF(411).INSEQ.NIN
      INTEGER RETMIN(14).RETSEC(14).NS
      REAL IRBAS.IRTIC.CFACT
      INTEGER IRDR.IPTR1.IPTR2.IPTR3.IIN.IOUT
      COMMON /IHDR/ INAM.ISCN.IDBUF.IDHDR.IDMIN.IRTMIN.
     1IRTSEC.INMBAS.IRBAS.IRTIC
      COMMON /INDATA/ INBUF.INSEQ.NIN
      COMMON /FLAGS/ EOF
      COMMON /UNITS/ IRDR.IPTR1.IPTR2.IPTR3.IIN.IOUT
      COMMON /COUEUE/ RETMIN.RETSEC.NS
C HAND THE BALL TO PADDOCK AND GET A RECORD
      CALL EPARD
      IEOF=EOF
      IF(EOF.LT.2)GO TO 10
      WRITE(IPTR1.20)
   20 FORMAT(1X.'***EXECUTION TERMINATING DUE TO IMPROPER EOF CONDITION
     1DETECTED IN RDSPC***')
      STOP
C CONTINUE WITH NORMAL PROCESSING
   10 IF(EOF.EQ.1)RETURN
C NORMAL SPECTRUM RETRIEVED - GET SCAN NUMBER
      SPCNO=ISCN
C CALCULATE CORRECTION FACTOR
      CFACT=IRBAS/999.0
C NOW SET UP AND LOOP SPEC ARRAY FROM INBUF
      DO 60 I=1.411
   60 AVGSPC(I)=CFACT*INBUF(I)
C UPDATE CIRCULAR QUEUE WHEN NECESSARY
      IF(IQUPD.EQ.0)RETURN
      NS=NS+1
      RETMIN(I14)=IRTMIN
      RETSEC(I14)=IRTSEC
C MISSION ACCOMPLISHED - I HOPE....
      RETURN
      END
      SUBROUTINE WRTSPC(SPCNO.NUSPEC.DUBLT.RATE.NPK.TICIN)
C THIS COMPRISES THE LAST OF SIX SUBROUTINES FOR THE EPA
C I/O INTERFACE TO THE CLEANUP PROGRAM. THE CLEANED UP SCAN
```

```
C RECORD IS REFORMATTED AND NECESSARY INFORMATION IS RETRIEVED
C FROM A CIRCULAR QUEUE. THE LOGIC EXACTLY PARALLELS THAT FOR
C THE RTI LINEAR ACQUISITION SYSTEM VERSION.
C                                                WM.F.HARGROVE
C                                                RTI 11/26/78
C
      INTEGER SPCNO.NUSPEC(411).DUBLT.NIN.RATE
      REAL TICIN
      INTEGER INAM(3).ISCN.IDBUF(2).IDHDR.IDMIN.IRTMIN.IRTSEC.I7.
     1INMBAS.INBUF(411).INSEQ.ONAM(3).OSCN.ODBUFF(2).ODHR.
     2ODMIN.ORTMIN.ORTSEC.ONMBAS.OTBUFF(411).OTSEQ.NOUT.NPK
      REAL IRBAS.IRTIC.ORBAS.ORTIC.CFACT
      INTEGER NS.RETMIN(14).RETSEC(14)
      COMMON /IHDR/ INAM.ISCN.IDBUF.IDHDR.IDMIN.IRTMIN.
     1IRTSEC.INMBAS.IRBAS.IRTIC
      COMMON /INDATA/ INBUF.INSEQ.NIN
      COMMON /CQUEUE/ RETMIN.RETSEC.NS
      COMMON /OHDR/ ONAM.OSCN.ODBUFF.ODHR.ODMIN.ORTMIN.
     2ORTSEC.ONMBAS.ORBAS.ORTIC
      COMMON /OTDATA/ OTBUFF.OTSEQ.NOUT
      COMMON /UNITS/ IRDR.IPTR1.IPTR2.IPTR3.IIN.IOUT
C NOW START FORMING THE HEADER RECORD
      DO 10 I=1.3
   10 ONAM(I)=INAM(I)
      OSCN=SPCNO
      DO 20 I=1.2
   20 ODBUFF(I)=IDBUF(I)
      ODHR=IDHDR
      ODMIN=IDMIN
C RETRIEVE SCAN RETENTION TIMES FROM CIRCULAR QUEUE
      I7=NS-7-14*((NS-8)/14)
      ORTMIN=RETMIN(I7)
      ORTSEC=RETSEC(I7)
C FORM OUTPUT ARPAY AND FIND BASE PEAK MASS AND INTENSITY
C TRANSFER AND DETERMINE NUMBER OF PEAKS AND BASE PEAK
      ORBAS=0
      ONMBAS=0
      ORTIC=0.0
      NOUT=0
      DO 40 I=1.411
      IF(NUSPEC(I).LE.0)GO TO 30
      ORTIC=ORTIC+NUSPEC(I)
      NOUT=NOUT+1
   30 CONTINUE
      IF(NUSPEC(I).LT.ORBAS)GO TO 40
      ORBAS=NUSPEC(I)
      ONMBAS=I
   40 CONTINUE
      ONMBAS=ONMBAS+39
C CALCULATE CORRECTION FACTOR
      IF(ORBAS.GT.0)GO TO 50
      WRITE(IPTR2,45) SPCNO
   45 FORMAT('-',/'*** WARNING - 0 BASE PEAK IN WRTSPC. SCAN #'.I5.'***')
      RETURN
C CALCULATE CORRECTION FACTOR
   50 CFACT=999.0/ORBAS
C NORMALIZE SPECTRUM
      DO 60 I=1.411
   60 OTBUFF(I)=CFACT*NUSPEC(I)+.05
C NOW EVERYTHING IS READY FOR PADDOCK'S OUTPUT ROUTINE
      CALL EPAWRT
      RETURN
      END
```

APPENDIX B

"EXTRACTION OF MASS SPECTRA FREE OF BACKGROUND AND NEIGHBORING COMPONENT CONTRIBUTION FROM GAS CHROMATOGRAPHY MASS SPECTROMETRY DATA"

R. G. DROMEY, M. J. STEFIK, T. C. RINDFLEISCH, AND
A. M. DUFFIELD, ANAL. CHEM., 48, 1368(1976)

Reprinted with permission from Analytical Chemistry.
Copyright by the American Chemical Society.

Extraction of Mass Spectra Free of Background and Neighboring Component Contributions from Gas Chromatography/Mass Spectrometry Data

R. G. Dromey,[1] Mark J. Stefik, Thomas C. Rindfleisch,* and Alan M. Duffield[2]

Departments of Computer Science, Genetics, and Chemistry, Stanford University, Stanford, Calif. 94305

An effective, minicomputer-based method is described for systematically extracting resolved mass spectra of mixture components from GC/MS data. Using tabular peak models derived directly from the raw data, the spectra have column bleed background removed and are corrected for interference from neighboring elutants and peak saturation. Individual components are detected in the data by means of a pair of histograms which statistically characterize the positions of mass fragmentogram peak modes. These data-adaptive corrections avoid costly iterative numerical procedures and allow obtaining representative mass spectra from GC/MS data of complex mixtures on a routine basis. Using this approach, components that elute within less than two spectral scan times of each other can be detected and their mass spectra well resolved.

With the increasing application of gas chromatography/mass spectrometry (GC/MS) systems to mixture component identification in biomedical research (*1*, *2*) and other areas (*3*), it has become important to be able to systematically isolate and identify minor components in the complex mixtures being analyzed. Because of instrumentation limitations, the mass

---

[1] Present address, Research School of Chemistry, Australian National University, Canberra, A.C.T., Australia.
[2] Present address, School of Physiology and Pharmacology, University of New South Wales, 2033, Australia.

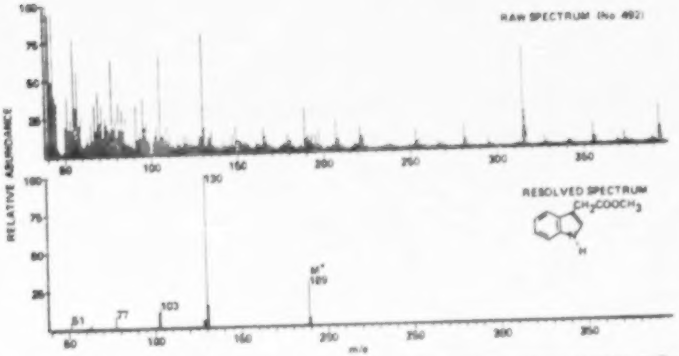


Figure 1. (a) Spectrum of indole acetic acid 3-methyl ester from a GC/MS analysis of human urine before processing. (b) Resolved spectrum of indole acetic acid 3-methyl ester

spectra obtained from a GC/MS analysis of a complex mixture are often markedly different from the spectra of the corresponding pure compounds. Differences may be caused by contributions from unresolved neighboring components during partial separation and also from GC septum and column bleed. These extraneous contributions may severely distort the relative intensities of ions in the mass spectrum of a particular component as well as contribute peaks that are not characteristic of the component being examined. Characterization and removal of these spurious ion contributions is especially important in the analysis of minor constituents where the mass spectra of interest may be substantially masked or distorted.

Our objective has been to implement a solution to these problems which is general and can systematically and reliably resolve GC/MS data with a minimum of human intervention. At the same time we have constrained the design so that the programs can run on a laboratory minicomputer. The first of these objectives has necessitated the use of a relatively complex mathematical treatment of the GC peak profile analyses as compared to that previously reported by Biller and Biemann (*4*). Both the present approach and that of Ref. 4 are based on analyses of mass fragmentogram profiles (*4–6*), a method which has been in use in various laboratories including our own, for a number of years. The method described here, however, differs substantially in the extraction of information from the profiles and thereby avoids several serious limitations inherent in the system described previously (*4*). By using tabular models of the elutant peak shapes together with a polynomial approximation to the GC background, and by deploying the elutant location and multiplicity information gained in analyzing individual fragmentogram profiles to assist in analyzing the others, we can achieve significant advantages in the quality of the reduced data. These include better final GC resolution, the proper assignment of ions to resolved elutant spectra (whether or not they are shared between neighboring components), more accurate spectral amplitudes free from background contributions, and the recovery of usable information from distorted data as in saturated peaks. We feel these improvements are important to a system which can reliably extract component spectra of sufficiently high quality from GC/MS runs to enable more definitive library matching, easier human interpretation of unknowns, and even the addition of extracted spectra to a library as authentic spectra. In our experience, these are essential assets for a GC/MS data system which is to be routinely applied in medical research and amply justify the complexity of the analysis.

EXPERIMENTAL

The GC/MS computer system used in this investigation consists of a Finnigan 1015 Quadrupole mass spectrometer interfaced to a PDP-11/20 minicomputer system for data acquisition. In one frequent mode of operation a complete mass scan (from mass 40 to 450) is completed each 3.7 s and 600 consecutive mass spectra are collected during a typical GC/MS analysis. Our initial experience of comparing the experimental mass spectra from a complex GC/MS analysis with a library of known mass spectra produced very poor results because of contamination of the experimental data by spectra of column bleed and of neighboring, unresolved components. Tolerable matches were only achieved when a component was present in large quantity in the GC/MS analysis. In order to overcome these problems, we have developed a computer program capable of systematically extracting from the raw GC/MS data, spectra representative of the pure elutant compounds.

The raw mass spectrum (Figure 1a) of indole acetic acid 3-methyl ester obtained from a GC/MS analysis of the acidic fraction (after methylation) of human urine typifies this situation. This component elutes at or near spectrum number 492 in the total ion plot (TIC) shown in Figure 2. Closer examination of Figure 2 shows that this component is submerged both in mass spectral contributions from neighboring components and background from GC column bleed. For comparison, a library spectrum (*7*) of indole acetic acid 3-methyl ester is shown in Figure 3. Figure 1b shows how, after processing the raw GC/MS data by the method described below, we can retrieve a high quality mass spectrum of indole acetic acid 3-methyl ester free from the environmental perturbations present in Figure 1a.

Thus, in the systematic analysis of GC/MS data the problem is first to detect where in the GC trace each component shows its maximum ion intensity and then to extract from these regions representative spectra for each of the detected components. The extracted mass spectra should be as free as possible from intensity distortions relative to their library counterparts, and from the presence of extraneous ions (e.g., peaks from either neighboring components or gas chromatographic column bleed).

DESCRIPTION OF METHOD

To obtain a reliable solution to these problems, it is necessary to analyze a number of spectra on either side of the ion current maximum for each elutant. A basic assumption of our approach is that the mass spectra of two neighboring unresolved elutants can be distinguished; that is, there exist some masses for which ions occur in the mass spectrum of one component but not in the other and vice versa. A schematic representation for two closely spaced elutants is given in Figure 4. By locating the "resolved" or singlet fragmentogram peaks at such masses (detected on the basis of profile mor-

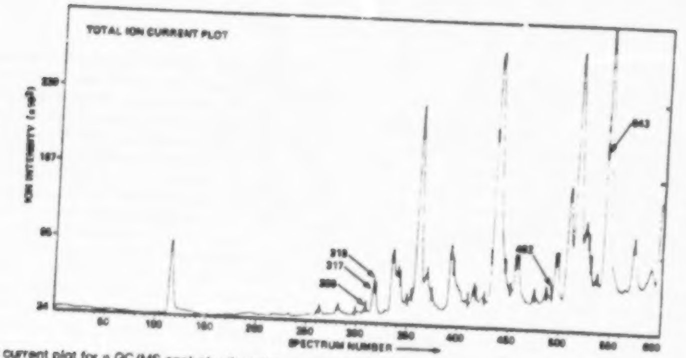


Figure 2. Total ion current plot for a GC/MS analysis of a urine sample. Components were found at vertical bar marks on TIC

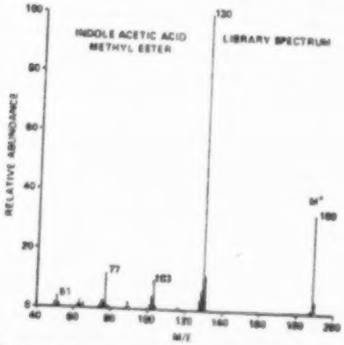


Figure 3. Mass spectrum of indole acetic acid 3-methyl ester taken from a library of biological compounds

phology), one can infer directly the positions of the elutants present and derive tabular models of the individual peak shapes. These models can be used subsequently to separate the unresolved fragmentogram complexes. The use of tabular peak models derived from the data itself accurately accommodates the a priori unknown peak profiles of particular elutants without solving for multiparameter, nonlinear model functions. Since the data are sampled often enough to satisfy the sampling theorem (*8*), these tabular models contain the necessary information to reconstruct a continuous peak envelope and can therefore be used as if they were continuous analytical models. For the typical peak shapes encountered, the collection of 5–10 mass spectra per singlet elutant peak represents a sampling frequency greater than twice the Fourier bandwidth of the peak. In addition, the mass by mass analysis of the fragmentogram peak complexes facilitates the mass dependent subtraction of background. (The large variation in background levels for different masses is a function of both the type of GC column used and the mixture being analyzed.)

By addressing the problem in this way, we have been able to produce accurate intensity information for the processed mass spectra and simultaneously distinguish with greater confidence which masses contribute to particular elutant spectra. We have been able to distinguish reliably elutants coming off within one and a half to two spectral scan times of each other. The succeeding sections discuss in more detail the procedures used to detect and resolve the mass spectra of unique elutants.

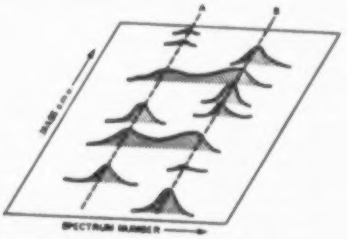


Figure 4. A schematic representation of the set of partial mass fragmentograms for two closely spaced elutants. Components A and B have some masses in common

Detection of Elutants in GC/MS Data. Elutant detection involves finding the location of each mixture component in the GC/MS data, even if it does not have a corresponding peak maximum in the overall total ion current trace. Ideally for a given elutant, the fragmentograms for all its ion masses will show maxima at the same time and, in practice, this holds for well-resolved materials. However, for partially resolved mixtures, the complicating factors of peak overlap and background contributions can caus fragmentogram maxima for neighboring components to show significant variation in their positions on the time axis. Reliable position information for each elutant is best derived from the fragmentogram profiles containing singlet peaks for that elutant, that is, from fragmentograms at those spectral masses unique to the elutant relative to its neighbors.

The approach used for elutant detection is to compute two histograms of candidate singlet peak positions and to select as elutant locations significant histogram maxima. The first histogram measures the number of singlet mass fragmentogram profiles which reach maxima in each time interval. The second histogram measures the total singlet ion intensity above background at these maxima. These two types of histogram contribute complementary information for judging elutant locations. At a given elution time, the histograms include fragmentogram peak maxima from all masses over seven spectra. The position of each maximum is determined by a parabolic least squares interpolation about the top five points in the sampled peak data. If the intensities of the five points contributing to the maximum are Y_{-2}, Y_{-1}, Y_0, Y_1, and Y_2, then the expression for the time coordinate of the maximum is

$$t = \frac{7(2Y_{-2} + Y_{-1} - Y_1 - 2Y_2)}{10(2Y_{-2} - Y_{-1} - 2Y_0 - Y_1 + 2Y_2)}$$

The time coordinates of maxima are estimated to one third of the time to collect each spectrum in order to separate very close neighbors. Because we measure peak locations to one third of a spectral scan time, appropriate shifts are also included to account for the fact that higher masses are measured later in each spectral scan than lower masses. To build these histograms, the program examines the profiles of each mass fragmentogram in the data. Only peaks with intensities above a prescribed threshold are added into their appropriate time positions in the histograms. Peaks that are obvious multiplets (multiple extrema) are not incorporated into the histograms but are marked for later resolution. After all of the histogram information is collected for a given region, components are defined to be detected at locations where both the intensity and peak count histograms show maxima that are above a threshold. This statistical approach, looking for "clusters" of fragmentogram peaks in the histograms, does not depend upon a correct decision for each peak but rather on a preponderance of good decisions looking over all of the data. It will fail to resolve elutants very close together which do not have enough distinguishing mass spectral components as described above. In general, however, using this approach we are able to detect and resolve spectra reliably that elute with a separation in time as small as one and a half to two spectral scan times. (Two scan times correspond to 25% of a typical GC peak width at the scan rate we use.) Elutants this close often do not show multiple extrema in the fragmentogram profiles of masses common to both spectra and could not be separated properly except for this type of procedure. If two elutants are separated by less than 1.5 scan times, resolution becomes less certain depending on their relative concentrations and mass spectral distinctness.

Estimation of Spectral Intensities and Background for Well Resolved Elutants. Once the locations of elutants in the GC effluent have been determined, we proceed to compute a resolved spectrum for each material. To illustrate the principles involved in spectral amplitude and background estimation, we consider the simple case of an elutant that is well separated in time from its nearest neighbors. This analysis will be extended to the more complicated case of multiplet resolution in a later section. By "well-separated" we imply only that there are no maxima in the elutant detection histograms for three or four spectra on either side of the elutant under consideration. In such a situation, each of the mass fragmentogram profiles in the vicinity of the elutant will consist of a background on which is superimposed a peak with amplitude representative of the elutant spectral component at that mass. The background (contributed by both GC column bleed and possible tailing from nearby, high-concentration elutants) is distinguished from the elutant peak by the fact that it varies much more slowly with time. Reasonable estimates can be made by assuming that, for any particular mass fragmentogram, the background amplitude varies at most linearly with elution time in the vicinity of a given elutant. This approach to background determination, using the actual fragmentogram characteristics around each elutant, automatically tracks changes in the bleed levels observed during a run. It should be noted that our model is a first-order approximation subject to some error. A more accurate approximation would involve representing the background variations over a larger span of spectral scans than we are able to manage with the current program organization and computer memory limitations. We feel that the linear estimate is justified, however, in that it produces results within the error limits from other data uncertainties.

To complete the estimation process, we use a model peak to determine the contribution of each mass fragmentogram to the elutant spectrum. Much work has been done on the analytic approximation of gas chromatographic peak shapes (*9, 10*). Our experience has been that relatively simple models do not adequately approximate the range of shapes encountered and more complex models require large amounts of computing to determine model parameters. Noting that a separate model must be developed for each elutant and with a view toward obtaining the peak shape and definition necessary for multiplet resolution within reasonable computing resources, we have approached the problem by using tabular peak models taken from the data itself. Such models, defined at discrete sample points, can be evaluated at any required intermediate point by interpolation (since the sampling theorem is satisfied) and automatically reflect any peak asymmetries which may be present. For a given elutant, the model will be independent of mass, assuming that relative molecular fragmentation probabilities do not change with elutant pressure within the mass spectrometer. A number of criteria should be satisfied by the tabulated model peaks. They should be singlet peaks superimposed on as small a background as possible and they should be relatively intense in order to ensure a good signal-to-noise ratio and good definition of peak skirts.

Candidate singlet peaks may be distinguished from doublet or background peaks by the feature that they are relatively sharp. One way to measure peak "sharpness" is to use a logarithmic rate function defined as follows:

$$\text{rate} = \sum_{i=1}^{3} \left[\frac{(Y_{i-1} - Y_i)}{Y_i} + \frac{(Y_{-(i-1)} - Y_{-i})}{Y_{-i}} \right]$$

where the Y_i are evaluated at equal scan widths at each side of the mode of the peak. It can be seen that this rate will be large for peaks which are sharp and smaller for peaks which are broad. The rate as defined is also independent of amplitude for peaks of identical shape. A peak with a computed rate below a threshold appropriate to the experimental conditions is considered to be either an artifact of the gas chromatograph (background peak) or a multiplet and is not included in the detection histograms.

During the process of computing the detection histograms, a list is kept of the unimodal fragmentogram peaks having the highest *rate* factors in the region under analysis. When a component is detected in a given region, a model peak is then immediately in hand that can be used in the peak height estimation and background removal process. The local minima just on either side of the model peak are used as estimates of the local background (a straight line through the greatest of these minima is removed before the model peak is used for analysis). The selection of the peak with the highest rate factor as our model peak has worked well in producing models which are singlets and suffer least from interference by background and neighboring fragmentogram peaks.

Given the fragmentogram peak model for this case of a well-separated elutant, we can now correct the individual mass fragmentograms for background and estimate true mass spectral intensities for the elutant. For the fragmentograms exhibiting peak maxima "near" the location of this elutant (see below for detailed selection criteria), each peak in the set is quadratically interpolated to align it on a common time origin (this removes the time shift between collection of low and high mass data). This is done by fitting a parabola through successive groups of three points near the peak mode and interpolating to give four equally spaced points about the mode, separated by one spectral scan time. With the peaks in this standard form, they are ready for the least squares analysis below. Assuming a linear background model over the region of 5 to 10 scan intervals under consideration, the local background B_t at time t is approximated by

$$B_t \sim c + dt$$

where c is the background offset and d is its slope. The interpolated elutant peak model is normalized to unit area and has amplitudes P_t at times t. Then for a given mass fragmentogram, the amplitude of the actual fragmentogram profile Y_t at time t can be approximated by

$$Y_t \sim pP_t + (c + dt)$$

where p measures the elutant amplitude above background. Note that this model assumes a superposition principle based on the earlier assumption of constant relative fragmentation probabilities and a linear encoding of ion current information. If ion current data are obtained from nonlinear electronic systems or read from film, the peak model itself would be amplitude dependent and this linear analysis could not be applied until appropriate amplitude linearization corrections were made. From the above model, we can derive a least squares estimate for the elutant amplitude p and the background parameters c and d by minimizing the error function

$$E = \sum (Y_t - pP_t - c - dt)^2$$

according to the conditions

$$\frac{\partial E}{\partial p} = \frac{\partial E}{\partial c} = \frac{\partial E}{\partial d} = 0$$

The summation in the error function is over all available points in the peak profile as well as the neighboring background points within the window of scans contained in the computer's memory. These conditions yield three linear equations in the three parameters which can be solved by standard techniques (*11*). From the solution of these equations for the value of p, we get the spectral intensity for the mass under consideration. This analysis is applied to all mass fragmentograms with maxima near the elutant location to obtain the complete, intensity-corrected spectrum. It is worth noting that this method, using a tabular model peak derived from the data and elutant locations obtained from the detection histogram analysis, reduces the calculation for each mass spectrum intensity to the solution of a set of linear equations. Specifically, this avoids iterative methods for determining the parameters of a theoretical peak model and for determining elutant time positions.

Fragmentograms are selected for this analysis on the basis of several criteria. Given the nominal elutant position from the detection histogram analysis, a fragmentogram is excluded (mass spectrum assigned zero intensity) if it has no local peak maximum or if its maximum is displaced from the reference elutant position by more than two thirds of a spectral scan time on either side. Each fragmentogram peak meeting this test must also have an acceptably high rate factor, to be included in the analysis. For peaks of masses greater than 200 amu, we require a rate factor greater than 25% of the rate for the model peak. This restriction is useful for eliminating contributions caused by peaking in column bleed components. In carefully examining GC/MS data sets, we can observe that masses characteristic of the spectra of column bleed components show maxima in their mass fragmentograms just prior (one to two spectra) to the elution of an actual component. In essence, the component appears to "push" the bleed out ahead of itself. Because these peaks are formed by a different process than normal elutant mass fragmentogram peaks, they usually have a much broader shape. Consequently their rate factors will be significantly reduced and they can be eliminated by the rate threshold criterion. The combination of the fragmentogram peak location criterion together with the minimum rate criterion effectively discriminates against extraneous contributions to the intensity-corrected spectra without removing authentic mass peaks.

Extraction of Poorly Separated Elutant Spectra. Many instances arise in the analysis of GC/MS data where two or more elutants are poorly resolved by the gas chromatograph. The resulting mass spectra in such a region exhibit ion intensity distortions which reflect the interactions (overlap) between adjacent elutants in addition to the ion contributions of background. The extension of the above procedures to the general case is not difficult. Through the histogram detection and model procedures, one can extract normalized peak models $P, Q, R, \ldots$ for the various elutants present. Then with the assumption of a linear background, the elutant contributions to each fragmentogram profile Y can be estimated by minimizing the error function

$$E = \sum (Y_t - pP_t - qQ_t - rR_t - \ldots - c - dt)^2$$

with respect to the elutant amplitudes $p, q, r, \ldots$, and the background coefficients. Sets of linear equations result for each mass to extract the resolved spectra. In practice, we have not implemented this full procedure beyond the doublet case. Through the following approximations, reasonable results are achievable within available minicomputer resources. Using the histogram method described earlier, neighboring elutants are handled with a "look ahead" procedure. That is, information about an elutant that has just been detected is stored and the detection algorithm is applied to the data in the immediate neighborhood by extending the range over which the detection histograms are calculated. If by including this extended region an additional elutant is detected, we record the position of its mode, select a model peak for this second elutant using the rate criterion, and initiate a doublet resolver algorithm. At present, the extended histograms project four spectral scan widths beyond the position where the first elutant of the multiplet was detected (limited by computer memory). The same criteria are applied as in the singlet case to decide which fragmentogram peaks belong to the pair of detected components. The model used to process the composite fragmentogram peaks (many of which may be singlets belonging to either elutant) assumes that there are two overlapping peaks superimposed on a linear background. The doublet model represents an oversimplification of some situations as, for example, in the case where 3 components elute within a very brief interval. By applying it, however, to successive pairs of elutant peaks (taking first-order account of peak tail contributions from any earlier elutant), it provides acceptable accuracy and peak resolution effectiveness. As indicated above, a fit of the two peak models P and Q with a linear background to the fragmentogram profile Y may be described by the approximation

$$Y_t \sim pP_t + qQ_t + c + dt$$

Minimizing an error function analogous to the earlier singlet case results in 4 linear equations in the peak amplitudes p and q and the background parameters c and d. This set of equations again can be solved by standard methods.

In cases where peaks are actually singlet peaks, the solution should yield zero for the amplitude of the missing component. In practice, for such cases the amplitude of the second component is a very small positive or negative value which is representative of how well the model fits the data. Amplitude results for masses that belong to the second component of the doublet are stored temporarily until this component is moved into the processing window at which time they are incorporated into the analysis of the newly detected component.

Reconstruction of Saturated Peaks in Elutant Spectra. From a practical viewpoint, a fairly common occurrence in

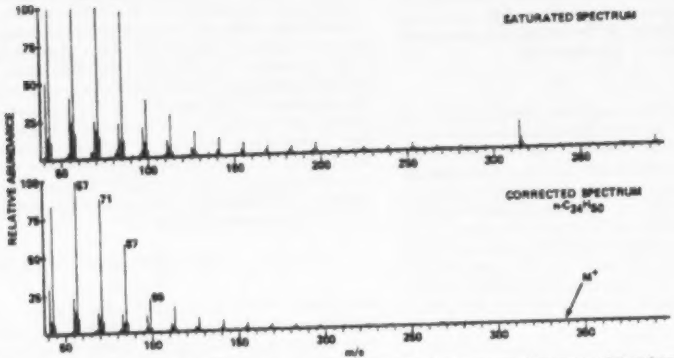


Figure 5. (a) Saturated unprocessed mass spectrum for tetracosane. (b) Spectrum of tetracosane after processing and correction for saturation

GC/MS data collection systems is the problem of mass peak saturation. Saturated peaks occur when the concentration of a component in the ion source is such that for one or more ion masses the detection system analog-to-digital converter becomes overloaded. Saturated peaks are easily detected because of their characteristic flat tops which have an amplitude determined by the overload limit of the detection system (e.g., the saturation value is 4095 for a twelve-bit analog-to-digital converter).

To obtain accurate amplitudes for component spectra that include saturated mass peaks, we must reconstruct these peaks to estimate their true amplitudes. A convenient way to do this in the singlet case is to use the least squares model that we derived in the preceding sections. To actually apply it for reconstruction of saturated mass fragmentogram peaks, we need to make a small modification to the equations. Instead of summing over all the points in the peak, we sum over only those points that are not saturated in the fragmentogram. As an estimate of the peak mode, we use the mode of the intensity histogram for the component being analyzed. An example of reconstruction of a mass spectrum with saturated ion intensities is given in Figure 5. Figure 5a shows a saturated spectrum of tetracosane (spectrum number 545 in Figure 2) and Figure 5b is the corresponding reconstructed spectrum. It is clear that the reconstructed spectrum will give a far better match with a library spectrum than the saturated spectrum which is badly saturated at masses 43, 57, and 71.

Before leaving the discussion of saturation, we should point out that we have not in practice extended the procedure for saturation correction of singlet peaks to the doublet case as we believe that it would be inadequate for reliable intensity estimates. If too many points are overloaded, there will be insufficient data to accurately estimate the amplitude of each multiplet component. Despite such correction algorithms, there is no substitute for the collection of good quality raw data at the start.

RESULTS AND DISCUSSION

The program based on the algorithm outlined in the preceding sections has been tested on a wide variety of biological samples. It fits comfortably into a DEC PDP 11/45 computer (with 28K words of memory) and takes approximately 8 min to analyze a raw GC/MS data set of 600 mass spectra (scanned from masses 40 to 450). Much of this time is spent in reading the raw data from the disk and other input–output operations. Copies of the program, which is written in FORTRAN, are available from the authors. Currently, this program forms part of an automated analysis system for the GC/MS analysis of urine and blood samples. The program reduces the raw GC/MS data set of approximately 600 spectra to a set of about 60 resolved elutant spectra which are then matched against a library of mass spectra of biological compounds. This whole process takes about 20 min and produces an analysis of the sample, with known compounds in the mixture identified and the remaining unknown set marked for further study by chemists or other DENDRAL programs (*12–14*).

In evaluating performance of the program, a major issue is how well it is able to detect elutants in the data. The vertical bars on the TIC (Figure 2) indicate all the places where the program detected and isolated a component from the raw GC/MS data. The program's power of detection is illustrated for example by the elutant detected near spectrum number 492 in the total ion current plot shown in Figure 2. Although there is no evidence of a maximum in the TIC in the region near 492, the program was able to detect and isolate a good quality spectrum of indole acetic acid methyl ester (Figure 1). In the raw data, this spectrum is clearly submerged in background and overlapping contributions. A comparison of the resolved spectrum (Figure 1b) with a library spectrum (Figure 3) shows that the basic spectral intensity profiles are very similar even including the very low intensity ion of mass 89. Some very small ions (of intensity less than 5% relative abundance) are absent from the resolved spectrum because they have been lost in the background noise. It is worth noting that there are no peaks present in the resolved spectrum that are not in the library spectrum, that is, the extraneous mass spectral peaks in the raw data including peaks at masses 105, 253, and 315 are not included in the resolved spectrum. The relative intensities of the mass spectral peaks at masses 51, 62, 65, and 77 have been changed significantly from their levels in the raw data. This illustrates the importance of correcting the intensities for background. The mass spectral peaks at masses 51, 52, 63, 78, and 129 appear to maximize near spectrum 496 in the raw data instead of spectrum 492 because of the overlapping contributions of a poorly resolved elutant. Similar examples of the power of this technique exist in other parts of the GC profile in Figure 2.

The detectability of unresolved elutants is clearly a function of their amplitude relative to neighboring components and background. One way to characterize this is to measure the ratio of the total ion intensity (sum of the mass spectrum amplitudes) in the resolved spectrum compared to that in the unprocessed spectrum including background and overlap effects. The mass spectrum of the processed component at

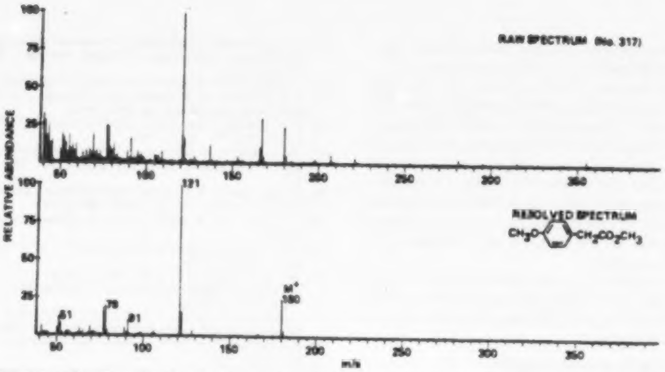


Figure 6. (a) Mass spectrum of 4-methoxyphenylacetic acid methyl ester before processing. (b) Resolved mass spectrum of 4-methoxyphenylacetic acid methyl ester

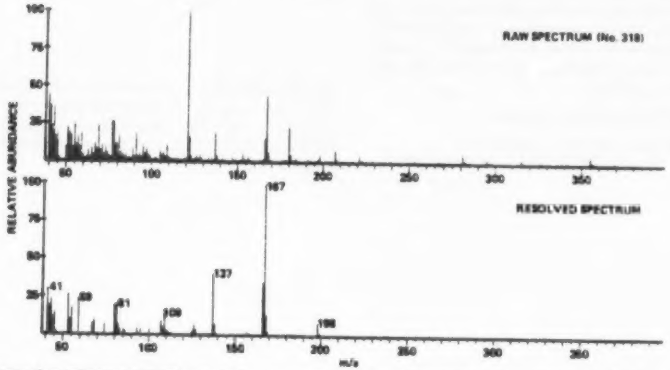


Figure 7. (a) Spectrum of an unknown aromatic ester before processing. (b) Resolved spectrum of unknown ester in Figure 7a

spectrum number 492 comprises only 4% of the total raw ion current. It can be expected that there will be problems detecting components with an ion current ratio that falls much below a level of 4%. Also if two compounds elute within less than 1.5 to 2 spectral scan times of one another, there is an increasing chance that the program will make the wrong decision as to whether there is one or actually two elutants present. Such errors are dependent on the ion current ratio between adjacent elutants, the similarity of their mass spectra, and the stability with which peak positions can be determined.

As an example of doublet resolution, consider the region near spectrum numbers 317 and 318 in Figure 2. The program detects that there are two elutants present and Figures 6 and 7 illustrate the raw and resolved spectra at these locations. The spectrum in Figure 6b is a good representation for 4-methoxyphenylacetic acid methyl ester. The other component is an unknown aromatic ester.

We have evaluated the efficiency of background removal for singlet elutants by examining their mass fragmentograms. After calculating the least squares peak and background levels, we concluded that the computed results are consistent (within 5–10%) with human estimates. They tended to be less accurate for very weak peaks whose shapes were more sensitive to noise distortions.

For the multiplet case, where the peak profiles can be considerably more complex, there is a stronger possibility that the model will not produce accurate amplitude information. In such cases, as when there are three rather than two elutants present, there is a danger that background contributions will be incorrectly estimated particularly with the limited number of scans that can be held in our minicomputer memory at one time. We feel however that use of a more complex model for triplets is not likely to be able to guarantee much greater precision. A sequential application of the doublet model has produced acceptable results in our experience with the program. Problems most frequently occur when a small amount of an elutant occurs just prior to, or just after, an elutant of high concentration. The intensities of peaks in the small elutant that are common to the large elutant tend to be less accurately calculated than singlet peaks and sometimes may even be discarded as negligible if their intensity relative to the large peak falls much below 10%. This may be especially important for the molecular ions of compounds with the same molecular weight which would be expected to elute near each other.

Comparison with library mass spectra has indicated that correction of intensities for saturated singlet peaks is satisfactory. However, as expected, the accuracy of the calculation decreases as peaks become more heavily saturated. In our case, we are working with model peaks that extend over nine points (i.e., nine scan widths). If more than four of a peak's nine points are saturated, we can expect that its estimated intensity will have only limited accuracy because there is insufficient information left to accurately characterize its shape.

Conditions arise in the raw GC/MS data for which it is not possible to extract resolved mass spectra unambiguously. One case is when the elutant-to-background ratio falls significantly below 5%. In these cases, the very weak intensity ions, including isotope ions, usually do not appear in the resolved mass spectra. The other difficulties arise when it is not possible to detect the presence of multiple elutants because they occur within less than one mass spectrum scan time of each other. In this case, the processed spectrum represents the mixture of the two elutants.

In general, we have found that the present system works very well and is capable of detecting and isolating high quality representative mass spectra in GC/MS experiments involving complex biological mixtures.

ACKNOWLEDGMENT

We thank D. Smith, W. Pereira, and W. Yeager who have contributed in a major way to the continued refinement of the computer programs implementing these algorithms and to the critique of results from their operational use in our laboratory. We also acknowledge the work of B. E. Blaisdell of Juniata College, Huntingdon, Pa., on an exploratory alternative approach to the present problem.

LITERATURE CITED

(1) R. A. Hites and K. Biemann, *Anal. Chem.*, **42**, 855 (1970).
(2) C. C. Sweeley, N. D. Young, J. F. Holland, and S. C. Gates, *J. Chromatogr.*, **99**, 507 (1974).
(3) W. H. McFadden, "Techniques of Combined Gas Chromatography/Mass Spectrometry: Applications in Organic Analysis", Wiley Interscience, London, 1973.
(4) J. E. Biller and K. Biemann, *Anal. Lett.*, **7**, 515 (1974).
(5) R. E. Summons, W. E. Pereira, W. E. Reynolds, T. C. Rindfleisch, and A. M. Duffield, *Anal. Chem.*, **46**, 582 (1974).
(6) R. N. Stillwell, 22nd Annual ASMS Conference—Mass Spectrometry, Philadelphia, Pa., 1974, p 454.
(7) "Mass Spectra of Compounds of Biological Interest", *U.S. At. Energy Comm. Rep.*, **No. TID-26553**, S. P. Markey, W. G. Urban, and S. P. Levine, Ed.
(8) See for example: R. S. Ledley, "Digital Computer and Control Engineering", McGraw-Hill Book Co., New York, 1960, p 742.
(9) E. Grushka, M. N. Myers, and J. C. Giddings, *Anal. Chem.*, **42**, 21 (1970).
(10) C. D. Scott, D. C. Chilcote, and W. W. Pitt, *Clin. Chem.* (*Winston-Salem, N.C.*), **16**, 637 (1970).
(11) See for example: S. D. Conte and C. de Boor, "Elementary Numerical Analysis: An Algorithmic Approach", McGraw-Hill Book Co., New York, 1972, p 241 and following.
(12) A. Buchs, A. B. Delfino, A. M. Duffield, C. Djerassi, B. G. Buchanan, E. A. Feigenbaum, and J. Lederberg, *Helv. Chim. Acta*, **53**, 1394 (1970).
(13) R. E. Carhart, D. H. Smith, H. Brown, and C. Djerassi, *J. Am. Chem. Soc.*, **97**, 5755 (1975).
(14) R. G. Dromey, B. G. Buchanan, D. H. Smith, J. Lederberg, and C. Djerassi, *J. Org. Chem.*, **40**, 770 (1975).

RECEIVED for review September 15, 1975. Accepted April 30, 1976. This work was supported by grants (Nos. RR-612 and GM-20832) from the National Institutes of Health and (No. NGR-05-020-632) from the National Aeronautics and Space Administration.